MARIA ALEJANDRA FERREIRA TORRES

**ANÁLISE GENÔMICA DE *STREPTOMYCES OLINDENSIS* DAUFPE 5622 E DE
SUAS VIAS CRÍPTICAS PARA A OBTENÇÃO DE NOVOS METABÓLITOS
SECUNDÁRIOS DE INTERESSE BIOTECNOLÓGICO**

Dissertação apresentada ao Programa de
Pós-Graduação em Microbiologia do
Instituto de Ciências Biomédicas da
Universidade de São Paulo, para
obtenção do título de Mestre em Ciências.

São Paulo
2015

MARIA ALEJANDRA FERREIRA TORRES

**ANÁLISE GENÔMICA DE *STREPTOMYCES OLINDENSIS* DAUFPE 5622 E DE
SUAS VIAS CRÍPTICAS PARA A OBTENÇÃO DE NOVOS METABÓLITOS
SECUNDÁRIOS DE INTERESSE BIOTECNOLÓGICO**

Dissertação apresentada ao Programa de
Pós-Graduação em Microbiologia do
Instituto de Ciências Biomédicas da
Universidade de São Paulo, para
obtenção do título de Mestre em Ciências.

Área de Concentração: Microbiologia

Orientador: Prof. Dr. Gabriel Padilla
Maldonado

Versão original

São Paulo
2015

DADOS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
Serviço de Biblioteca e Informação Biomédica do
Instituto de Ciências Biomédicas da Universidade de São Paulo

reprodução não autorizada pelo autor

Ferreira Torres, Maria Alejandra.

Análise genômica de *Streptomyces olindensis* DAUFPE 5622 e de suas vias crípticas para a obtenção de novos metabólitos secundários de interesse biotecnológico / Maria Alejandra Ferreira Torres. – São Paulo, 2015.

Orientador: Prof. Dr. Gabriel Padilla Maldonado.

Dissertação (Mestrado) – Universidade de São Paulo. Instituto de Ciências Biomédicas. Departamento de Microbiologia. Área de concentração: Microbiologia. Linha de pesquisa: Engenharia genética de *Streptomyces*.

Versão do título para o inglês: Analysis of *Streptomyces olindensis* DAUFPE 5622 genome and its cryptic pathways to obtain new secondary metabolites of biotechnological interest.

1. *Streptomyces* 2. Mineria de genomas 3. Anotação de genomas 4. Cluster biossintético 5. Cosmomicina D 6. Super-expressão de genes reguladores I. Maldonado, Prof. Dr. Gabriel Padilla II. Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação em Microbiologia III. Título.

ICB/SBIB0165/2015

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS

Candidato(a): Maria Alejandra Ferreira Torres.

Título da Dissertação: Análise genômica de *Streptomyces olindensis* DAUFPE 5622 e de suas vias crípticas para a obtenção de novos metabólitos secundários de interesse biotecnológico.

Orientador(a): Prof. Dr. Gabriel Padilla Maldonado.

A Comissão Julgadora dos trabalhos de Defesa da **Dissertação de Mestrado**,
em sessão pública realizada a/...../....., considerou

(☐) **Aprovado(a)**

(☐) **Reprovado(a)**

Examinador(a): Assinatura:
Nome:
Instituição:

Examinador(a): Assinatura:
Nome:
Instituição:

Presidente: Assinatura:
Nome:
Instituição:

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS

Cidade Universitária "Armando de Salles Oliveira"
Av. Prof. Lineu Prestes, 2415 – CEP. 05508-000 São Paulo, SP – Brasil
Telefone : (55) (11) 3091-7733 - telefax : (55) (11) 3091-8405
e-mail: cep@icb.usp.br

Comissão de Ética em Pesquisa

CERTIFICADO DE ISENÇÃO

Certificamos que o Protocolo CEP-ICB N° 696/14 referente ao projeto intitulado: “*Análise genômica de Streptomyces olindensis DAUFPE 5622 e expressão de vias crípticas para a obtenção de novos metabólitos secundários*” sob a responsabilidade de **Maria Alejandra Ferreira Torres**, foi analisado na presente data pela CEUA - COMISSÃO DE ÉTICA NO USO DE ANIMAIS e pela CEPSh- COMISSÃO DE ÉTICA EM PESQUISA COM SERES HUMANOS, tendo sido deliberado que o referido projeto não utilizará animais que estejam sob a égide da lei 11.794 de 8 de outubro de 2008, nem envolverá procedimentos regulados pela Resolução CONEP nº466 de 2012.

São Paulo, 23 de setembro de 2014.

PROF. DR. ANDERSON DE SÁ NUNES
Vice-Coordenador da CEUA - ICB/USP

PROF. DR. PAOLO M.A. ZANOTTO
Coordenador da CEPsh - ICB/USP

A mi familia que con tanto amor me han
permitido realizar cualquier locura!

AGRADECIMENTOS

A Dios porque en los momentos más difíciles y de desespero se acuerda de todos incluso nosotros los descarriados!

A mi mama por creer siempre que soy buena en todo, por darme animo en cada paso y por su amor incondicional.

A mi hermana por la paciencia en el día a día, por estar en cada recuerdo y por la mejor compañía.

A mi papa por su apoyo y sus consejos.

Al Prof. Dr. Gabriel Padilla por las oportunidades, los consejos, los ánimos y su apoyo inagotable.

Al Dr. Leandro Maza Garrido por su valiosa orientación, su amistad incondicional y su gran conocimiento que hicieron de este proyecto una divertida experiencia.

A la Dr. Manuella Nóbrega Dourado-Ribeiro por su amistad, guía intelectual, apoyo y animo durante este proceso de convertirme en Mestre.

A León Rueda, mi novio, quien valientemente decidió afrontar conmigo estos años de afanes, distancia, desespero y aprendizaje.

A los Alpinitos originales Ruth, María Paula, Liz, Maira, Juan y Roger por ser mi familia en este país, por el internet, las noches de estudio, las risas, los bandeijões y la locura.

A los brasileiros del Lab 164 (Simone, Mariana, Fernanda, Aline) que con paciencia y amistad me han acompañado en estos años de español, locuras, crisis y estudio.

Al Prof. Dr. Welington Luiz de Araújo e compañeros del Lab 136 que siempre nos ayudaron, nos acogieron, y sufrieron junto con nosotros, siempre dispuestos a ayudar.

Al Lado B, mis amigos entrañables de todas las nacionalidades que me acogieron desde el intercambio y hasta el día de hoy siguen apoyándome para que Brasil sea cada vez más mi segundo hogar.

A CAPES (Coordenação de Aperfeiçoamento Pessoal de Nível Superior) por el pequeño detalle de la bolsa.

A los que ya no están por siempre cuidar desde arriba y demás amigos y familiares que siempre estuvieron dispuestos a colaborar y a creer en mí.

"A man's got to do what a man's got to do. A woman must do what he can't"

Rhonda Hansome

"You cannot create experience, you must undergo it"

Albert Camus

RESUMO

FERREIRA-TORRES, M. A. **Análise genômica de *Streptomyces olindensis* DAUFPE 5622 e de suas vias crípticas para a obtenção de novos metabólitos secundários de interesse biotecnológico.** 2015. 103 f. Dissertação (Mestrado em Microbiologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2015.

A caracterização de novos clusters biossintéticos tem se tornado uma prioridade dentro da pesquisa com actinobactérias devido ao potencial biotecnológico para produzir novas moléculas com possível uso na indústria e farmácia. Estes produtos naturais têm readquirido interesse pelas características particulares que possuem, por exemplo: biodisponibilidade, especificidade de alvo e diversidade de estruturas químicas. Ademais, a maioria das vias biossintéticas para a produção destas moléculas, descobertas no genoma de bactérias do gênero *Streptomyces*, permanecem crípticas ou silenciadas em condições de cultura no laboratório. Isto tem levado ao desenvolvimento de diversas técnicas e metodologias para a expressão destas vias metabólicas. Uma destas estratégias é a super-expressão de genes reguladores específicos dentro destas vias, como as proteínas reguladoras de antibióticos em *Streptomyces* (SARP, pelas siglas em Inglês). O laboratório de Bio-Produtos no Departamento de Microbiologia no Instituto de Ciências Biomédicas, na Universidade de São Paulo, tem trabalhado com a linhagem *Streptomyces olindensis* DAUFPE 5622, que produz uma molécula do grupo das antraciclinas chamada Cosmomicina D. Esta possui uma alta atividade antitumoral e tem sido de grande interesse devido ao padrão de glicosilação. Utilizando sequenciamento *Shotgun*, foram obtidos 120 *contigs* representando o genoma de *S. olindensis*, os quais foram anotados e submetidos ao NCBI sob o número de acesso JJOH00000000. Utilizando o software antiSMASH e anotação manual, foram identificados trinta e três clusters envolvidos na produção de metabólitos secundários. Dentro destes, foram encontrados clusters gênicos para a produção de metabólitos previamente descritos, com diversas atividades biológicas, como a cosmomicina D, melanina, pigmento de esporos, fator-A, geosmina, albaflavenona, hopanoides, amiquelina, hidroxietoxina, desferrioxamina e fosacetamicina. Os outros vinte dois clusters possivelmente produzem derivados de compostos já conhecidos ou moléculas novas, motivo pelo qual, foi realizada uma análise de genômica comparativa para identificar, caracterizar e anotar cada via biossintética achada em *S. olindensis*. Subsequentemente, foram encontrados diversos reguladores específicos de vias metabólicas, tomando especial interesse os SARP. Foram escolhidos dois clusters biossintéticos para a super-expressão (Aminociclitol e um Policetídeo Tipo I) destes genes, levando a detecção do composto sob condições de cultura. Estudos posteriores permitirão identificar a estrutura química destes compostos e propor diferentes metodologias para a obtenção de produtos naturais de interesse biotecnológico, sintetizados por *Streptomyces olindensis*.

Palavras-chave: *Streptomyces*. *Genome mining*. Anotação de genoma. *Cluster* biossintético. Cosmomicina D. Super-expressão de genes reguladores

ABSTRACT

FERREIRA-TORRES M.A. **Analysis of *Streptomyces olindensis* DAUFPE 5622 genome and its cryptic pathways to obtain new secondary metabolites of biotechnological interest.** 2015. 103 p. Master thesis (Microbiology) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2015.

Characterization of new metabolic biosynthetic clusters has become a priority among researchers in actinobacteria because of their biotechnological potential to produce new molecules with possible use in the pharmaceutical and industrial field. These natural products regain interest because of their particular characteristics, for example: bioavailability, target specificity and chemical structure diversity. Moreover, most of the biosynthetic pathways for the production of these molecules that have been discovered on the genomes of *Streptomyces* remain cryptic or silenced under standard laboratory conditions, which lead to the development of different technical approaches to express these new clusters. One of the strategies is the overexpression of the specific regulatory genes within these pathways, such as *Streptomyces* Antibiotic Regulatory Protein (SARP) to detect these new products in different culture media. The Bio products Laboratory, in the Microbiology Department in the Biomedical Science Institute at São Paulo's University, has been working with *Streptomyces olindensis* DAUFPE 5622, the producer of the anthracycline Cosmomycin D, which has an important antitumoral activity and has attracted interest because of its distinctive glycosylation pattern. Using Shotgun sequencing the genome was obtained, organized in 120 contigs, annotated and submitted in the NCBI under the accession number JJOH00000000, and by further prediction were identified thirty-three secondary metabolite clusters using the antiSMASH server and manual annotation. Among them, were found gene clusters for the production of known metabolites with different biological activity such as Cosmomycin D, Melanin, Spore pigment, A-factor, Geosmin, Albaflavenone, Hopene, Amychelin, Hidroxyectoine, Desferrioxamine B and Fosacetamycin. The other twenty-two clusters were likely to produce derivatives of known compound or new molecules, which is why, a comparative genomic study was carried out to identify, characterize and annotate every genome cluster found in *S. olindensis*. Subsequently, several pathway-specific regulators in the described biosynthetic clusters were found, arising special interest the SARP, choosing two different pathways (Aminocyclitol and Polyketide Type I), to over express the regulatory genes, in order to detect the biological product under culture conditions. Further studies will allow the identity of the chemical structure of these products to be found and propose different approaches to obtain other natural products of interest synthesized by *Streptomyces olindensis*.

Keywords: *Streptomyces*. Genome mining. Genome annotation. Biosynthetic Cluster. Cosmomycin D. Overexpression of regulator genes.

LISTA DE ILUSTRAÇÕES

Figura 1 Ciclo de vida e características macroscópicas e microscópicas de <i>Streptomyces</i> sp.....	19
Figura 2 Diversidade estrutural de moléculas policetídeas	23
Figura 3 Módulos enzimáticos para a síntese de peptídeos não ribossomais pela condensação de aminoácidos	24
Figura 4 Via biossintética geral para os peptídeos ribossomais (RiPPs).	26
Figura 5 Estruturas dos terpenos mais produzidos por bactérias.....	27
Figura 6 Via biossintética geral para as moléculas derivadas do aminociclitol.	29
Figura 7 Diversidade estrutural e biológica dos fosfonatos produzidos por bactérias.	30
Figura 8 Síntese do Aminoetilfosfonato, precursor dos compostos C-P.	30
Figura 9 (A) Estrutura da Retamicina detectada por Bieber et al (1989), (B) Estrutura da Cosmomicina D detectada por Furlan et al (2004)	34
Figura 10 Agrupação gênica para a biossíntese da Cosmomicina D em <i>Streptomyces olindensis</i> DAUFPE 5622	35
Figura 11 Via biossintética da cosmomicina D em <i>Streptomyces olindensis</i> DAUFPE 5622	36
Figura 12 Classificação funcional dos CDS em <i>Streptomyces olindensis</i> DAUFPE 5622	47
Figura 13 Diagrama de Venn com os grupos de ortólogos da base de dados KEGG de acordo com a informação de vias metabólicas de <i>S. olindensis</i> , comparados com os grupos de ortólogos de quatro espécies de <i>Streptomyces</i>	51
Figura 14 Distribuição funcional dos 108 grupos de ortólogos únicos em <i>Streptomyces olindensis</i>	52
Figura 15 Árvore filogenética rDNA 16S de <i>Streptomyces olindensis</i> DAUFPE 562.. ..	53
Figura 16 Árvore filogenética concatenada de <i>Streptomyces olindensis</i> DAUFPE 5622 e diversidade de clusters de metabólitos secundários	54
Figura 17 Genes biossintéticos para a produção de γ -butirolactonas presentes em <i>S. olindensis</i> DAUFPE 5622	58
Figura 18 <i>Cluster</i> gênico para produção de hidroxiectoína presente em <i>Streptomyces olindensis</i>	59

Figura 19 <i>Cluster</i> gênico para produção de melanina presente em <i>Streptomyces olindensis</i>	61
Figura 20 <i>Cluster</i> gênico para produção de pigmento associado aos esporos presente em <i>Streptomyces olindensis</i>	63
Figura 21 Agrupação gênica para a produção da molécula da família dos fosfonatos	64
Figura 22 <i>Clusters</i> biossintéticos para a produção de geosmina, 2-metilisoborneol, albaflavenona e hopanóides.....	67
Figura 23 <i>Clusters</i> biossintéticos para a produção de terpenos sem produto conhecido em <i>S. olindensis</i>	68
Figura 24 <i>Cluster</i> gênico e sequência do peptídeo precursor do lantipeptídeo classe III	70
Figura 25 Organização do <i>cluster</i> gênico envolvido na biossíntese de um peptídeo não ribossomal (NRPS).....	72
Figura 26 Organização do <i>cluster</i> gênico envolvido na biossíntese de uma molécula da família dos aminociclitois.....	74
Figura 27 Organização do <i>cluster</i> gênico envolvido na biossíntese de um policetídeo Tipo I (Poliéter- <i>like</i>)	77
Figura 28 A). Padronização das condições de amplificação utilizando uma gradiente de temperatura.	80

LISTA DE TABELAS

Tabela 1 Metabólitos terpenóides amplamente estudados.	27
Tabela 2 Produtos naturais da família dos aminociclitóis.....	28
Tabela 3 Características gerais do genoma de <i>S. olindensis</i> DAUFPE 5622 comparado com outras espécies de <i>Streptomyces</i>	43
Tabela 4 Genes associados à replicação dos telómeros no genoma de <i>S. olindensis</i> DAUFPE 5622.....	45
Tabela 5 Genes associados à síntese de integrases e transposases.....	45
Tabela 6 Funções das proteínas na categoria de biossíntese, transporte e catabolismo de metabólitos secundários.....	48
Tabela 7 <i>Clusters</i> biossintéticos detectados pelo software antiSMASH dentro do genoma de <i>Streptomyces olindensis</i>	56
Tabela 8 Funções preditas das ORFs no cluster biossintético de hidroxiectoína.....	60
Tabela 9 Funções preditas das ORFs no cluster biossintético da melanina	62
Tabela 10 Funções preditas das ORFs no <i>cluster</i> biossintético para produção de pigmento de esporos.....	63
Tabela 11 ORFs preditas no <i>cluster</i> biossintético de fosfonato.....	65
Tabela 12 Funções preditas das ORFs no cluster biossintético para produção do antibiótico Lantipeptídeo	70
Tabela 13 Funções preditas do <i>cluster</i> biossintético para a produção do peptídeo não ribossomal (NRPS) Amiquelina-like	72
Tabela 14 ORFs preditas no <i>cluster</i> biossintético de aminociclitol.....	75
Tabela 15 Funções preditas do <i>cluster</i> biossintético para a produção do Policetídeo Tipo I – Poliéter-like.....	78

LISTA DE ABREVIATURAS E SIGLAS

aa	Aminoácido
ACP	Proteína transportadora de grupos acilo
AT	Aciltransferase
ATP	Adenosina 5'-trifosfato
CLF	Fator alongador da cadeia
CoA	Coenzima A
DNA	Ácido desoxirribonucleico
DH	Desidratase
DMSO	Dimetil-sulfóxido
ER	Enoil-redutase
ICM	Modelo de contexto interpolado
Kb	Kilobase
KR	Cetoreductase
KS	Cetosintase
Mb	Megabase
MFS	Família de facilitadores de transporte
NRPS	Peptídeo não ribossomal Sintase
ORF	Marco de leitura aberto
pb	Pares de bases
PCR	Reação em cadeia da polimerase
PKS	Policetídeo sintase
RiPP	Peptídeo de síntese ribossomal
RPM	Revoluções por minuto
RNA	Ácido ribonucleico
SARP	Proteína reguladora de antibióticos em <i>Streptomyces</i>
TE	Tioesterase

SUMÁRIO

1 INTRODUÇÃO	16
2 REVISÃO DA LITERATURA	18
2.1 Actinobacteria: O gênero <i>Streptomyces</i>	18
2.1.1 Características gerais	18
2.1.2 Ecologia dos <i>Streptomyces</i>	19
2.1.3 Características do genoma de <i>Streptomyces</i> spp.	20
2.2 Metabólitos secundários	21
2.2.1 Aspectos gerais.....	21
2.2.2 Policetídeos PKS.....	22
2.2.3 Peptídeos não ribossomais NRPS.....	24
2.2.4 Peptídeos Ribossomais RiPPs.....	25
2.2.5 Terpenos	26
2.2.6 Aminociclitolis.....	28
2.2.7 Fosfonatos	29
2.3 <i>Streptomyces</i> como fábrica de produtos naturais	31
2.4 Regulação de vias biossintéticas de metabólitos secundários e ativação de vias crípticas.....	32
2.5 Potencial biotecnológico de <i>Streptomyces olindensis</i> DAUFPE 5622	33
2.5.1 Biossíntese da Cosmomicina D em <i>Streptomyces olindensis</i> DAUFPE 5622.....	34
3 MATERIAIS E MÉTODOS	37
3.1 Análise genômica <i>in-silico</i>	37
3.1.1 Anotação e análise do genoma de <i>Streptomyces olindensis</i>	37
3.1.2 Análise filogenética de <i>Streptomyces olindensis</i> DAUFPE 5622.....	38
3.1.3 Genômica comparativa entre espécies de <i>Streptomyces</i>	38
3.1.4 Identificação e anotação de clusters biossintéticos envolvidos na produção de metabólitos secundários.....	39
3.2 Super-expressão dos genes ativadores SARP nas vias biossintéticas de produtos policetídeos e do aminociclitol.....	39
3.2.1 Desenho dos iniciadores	39
3.2.2 Crescimento das linhagens utilizadas.....	39
3.2.3 Conservação das linhagens	40
3.2.4 Extração de DNA total.....	40

3.2.5 Reações de PCR.....	40
3.2.6 Transformação de <i>E. coli</i> e protoplastos de <i>S. olindensis</i>	41
4 RESULTADOS E DISCUSSÃO	42
4.1 Generalidades físicas do genoma de <i>Streptomyces olindensis</i> DAUFPE 5622	42
4.2 Estudo funcional do genoma de <i>Streptomyces olindensis</i> DAUFPE 5622 ..	46
4.3 Genômica comparativa entre espécies de <i>Streptomyces</i>	49
4.4 Análise Filogenética de <i>Streptomyces olindensis</i> DAUFPE 5622.....	52
4.4.1 Filogenia inferida utilizando o gene 16S rDNA	52
4.4.2 Filogenia concatenada e comparação de diversidade de metabólitos secundários.....	54
4.5 Identificação do potencial biotecnológico utilizando a abordagem “genome mining” e descrição dos clusters biossintéticos achados no genoma de <i>S. olindensis</i> DAUFPE 5622	55
4.5.1 Descrição dos genes associados à produção de γ -Butirolactona	57
4.5.2 Anotação da via biossintética para produção de ectoína.....	59
4.5.3 Vias biossintéticas para produção de pigmentos	60
4.5.4 Cluster gênico da biossíntese do fosfonato em <i>S. olindensis</i>	64
4.5.5 Biossíntese de Terpenos em <i>S. olindensis</i> DAUFPE 5622	65
4.5.6 Peptídeos Ribossomais RiPPs.....	68
4.5.7 Anotação da via metabólica do peptídeo não ribossomal NRPS.....	71
4.5.8 Cluster do aminociclitol.....	74
4.5.9 Policetídeos preditos em <i>S. olindensis</i>	76
4.6 Amplificação dos genes SARP	80
5 CONCLUSÕES	81
REFERÊNCIAS.....	83
APÊNDICE A – Informação suplementar.....	94
APÊNDICE B - Funções preditas dos clusters achados em <i>S. olindensis</i>	95

1 INTRODUÇÃO

Historicamente, os metabólitos secundários de origem microbiana têm proporcionado uma grande quantidade de moléculas de interesse industrial para uso farmacêutico e outros campos na biotecnologia. Até 2010, tinham sido descritos mais de 500.000 sendo bioativos o 20%. Desta porção bioativa, 35.000 são produzidos por micro-organismos e o 35% é produzida por actinobactérias, sendo o restante produzido por fungos e outras bactérias (BÉRDY, 2012). Embora o desenvolvimento de novos fármacos tenha diminuído na última década, o interesse por novas estruturas químicas se mantém aberto na indústria devido as propriedades intrínsecas dos produtos naturais como a biodisponibilidade das moléculas, a complexidade estrutural e a especificidade de alvo (HARVEY; EDRADA-EBEL; QUINN, 2015). A interdisciplinaridade na pesquisa tem proporcionado diversas ferramentas para a busca de novas biomoléculas, encontrando, assim, diferentes *clusters* biossintéticos no genoma de vários micro-organismos, incluindo actinobactérias (BENTLEY et al., 2002; BÉRDY, 2012). Pode-se detectar domínios conservados típicos de enzimas que produzem moléculas com atividade biológica, caracterizadas quimicamente como policetídeos (PKS), peptídeos não ribossomais (NRPS), terpenos, peptídeos de síntese ribossomal, sideróforos, etc. Estes descobrimentos indicam um grande potencial biotecnológico para obtenção de novos antibióticos, compostos antitumorais e moléculas com outra aplicação biotecnológica (BÉRDY, 2005, 2012; CHALLIS, 2014; NETT; IKEDA; MOORE, 2009).

Em *Streptomyces coelicolor*, a actinobactéria modelo, foram encontrados 29 *clusters* biossintéticos dentro do seu genoma dedicados à produção de metabólitos secundários de interesse industrial, e em *Streptomyces avermitilis* detectaram-se 37 *clusters* para a produção de novas biomoléculas (NETT; IKEDA; MOORE, 2009). Estes dados são um indício do potencial codificado no genoma das actinobactérias, sendo candidatas à busca de novos produtos metabólicos, inicialmente pelo uso de ferramentas de bioinformática para sua caracterização e expressão. A maioria destes novos *clusters*, encontrados em diversos microrganismos, não é detectada em condições de fermentação padrão em laboratório, por isso são chamados de *clusters* ou vias crípticas (BÉRDY, 2012; NETT; IKEDA; MOORE, 2009; OCHI;

HOSAKA, 2013). Este fato leva à busca de novas estratégias para a expressão destes compostos, como a diversificação de meios de cultura e condições de crescimento para a obtenção de novas moléculas (OSMAC approach) (BODE et al., 2002) ou a produção em hospedeiros heterólogos (BALTZ, 2010; STARCEVIC et al., 2012).

No laboratório de Bioprodutos vem se trabalhando com a bactéria *Streptomyces olindensis*, isolada da região de Olinda, no estado de Pernambuco, Brasil. Esta bactéria produz o agente antitumoral Cosmomicina D, amplamente estudada, já que a molécula possui um padrão de glicosilação pouco comum, que potencializa a sua atividade contra diversas linhas celulares tumorais. Além disso, possui uma maior quantidade de genes de resistência que ajudam na saída do composto e na sobrevivência da bactéria (FURLAN et al., 2004; GARRIDO et al., 2006).

A sequência do genoma foi descrita previamente por Rojas e colaboradores em 2014, para ganhar um maior entendimento sobre o cluster gênico da Cosmomicina D e expandir as possibilidades na área de biossíntese combinatória para, assim, modificar a atividade biológica da molécula. Neste estudo é apresentada uma análise completa do genoma como também a comparação com outras espécies de *Streptomyces* relevantes no campo dos compostos bioativos. Finalmente é revelado o potencial codificado para uma variedade de novos produtos naturais, além do agente antitumoral Cosmomicina D, o que levará ao desenvolvimento de diferentes estratégias para a expressão e obtenção de novas moléculas.

2 REVISÃO DA LITERATURA

2.1 Actinobactéria: O gênero *Streptomyces*

2.1.1 Características gerais

O phylum *Actinobacteria* é um dos maiores grupos taxonômicos atualmente reconhecidos dentro do domínio *Bacteria*. Além disso, é um dos mais diversos em termos de morfologia, distribuição ecológica, patogenicidade, tamanho do genoma e conteúdo G+C (ALAM et al., 2010; VENTURA et al., 2007). Esta diversidade também está representada fisiológica e metabolicamente, como na produção de enzimas extracelulares e uma ampla variedade de metabólitos secundários (CHANDRA; CHATER, 2014; VENTURA et al., 2007).

Streptomyces é um dos 120 gêneros na ordem *Actinomycetales* que pertence ao phylum *Actinobacteria* (HOPWOOD, 2006). Devido à morfologia deste gênero (crescimento apical com geração de hifas) (Figura 1) acreditava-se que estes micro-organismos poderiam ser intermediários entre fungos e bactérias, até princípios dos anos 50, quando foram realizadas provas bioquímicas para determinar a composição da parede celular, sensibilidade a agentes antibacterianos, e finalmente determinação da ausência de membrana nuclear demonstrando que pertenciam aos procariotas (HOPWOOD, 1999). Estas bactérias Gram-positivas geralmente possuem um alto conteúdo G+C no DNA sendo em alguns casos maior a 70%, o que as distingue de outras bactérias gram-positivas como *Staphylococcus* ou *Bacillus* que apresentam um conteúdo de G+C de 50% (CHANDRA; CHATER, 2014; HOPWOOD, 2006).

O ciclo de vida de *Streptomyces* começa com a germinação do esporo em níveis adequados de nutrientes, e o crescimento apical do tubo germinativo dando origem ao micélio vegetativo que é composto por hifas ramificadas. Estas estruturas ao longo do tempo dividem-se para formar esporos durante um processo de diferenciação celular constituindo o micélio aéreo que é observado em condições de depleção de nutrientes, condições ambientais adversas e outros sinais no meio de cultura (Figura 1). Nesta etapa do ciclo de vida (fase estacionária) geralmente ativam-se as vias metabólicas associadas à produção de metabólitos secundários (CHANDRA; CHATER, 2014; HOPWOOD, 2006).

Figura 1 Ciclo de vida e características macroscópicas e microscópicas de *Streptomyces* sp.

Fonte: (SEIPKE; KALTENPOTH; HUTCHINGS, 2012)

2.1.2 Ecologia dos *Streptomyces*

Estas bactérias filamentosas encontram-se distribuídas amplamente no solo já que possuem a capacidade de colonizar a rizosfera utilizando o arsenal metabólico (CHANDRA; CHATER, 2014), onde desempenham um papel importante na decomposição de matéria orgânica e a reciclagem de nutrientes (GOODFELLOW; WILLIAMS, 1983; VENTURA et al., 2007).

Em ecossistemas de água doce e ambientes marinhos, foram descritos isolamentos de *Streptomyces* em vida livre ou associada a diversos organismos, como esponjas marinhas e caracóis, tornando-se de grande importância para a descoberta de novos produtos naturais (ALAM et al., 2010; OLANO; MÉNDEZ; SALAS, 2009; SEIPKE; KALTENPOTH; HUTCHINGS, 2012; UDWARY et al., 2007). Estes microrganismos também estão presentes nas raízes de plantas como simbios, atuando tanto como promotores de crescimento vegetal em espécies como o trigo (PGPR), ou como patógenos de tomate e tubérculos, como a batata (SEIPKE; KALTENPOTH; HUTCHINGS, 2012).

Em organismos invertebrados como vespas solitárias, formigas atíneas, formigas do gênero *Allomerus* e besouros, têm sido identificadas relações simbióticas com bactérias deste gênero, como mecanismo de defesa contra fungos patogênicos, no alimento das larvas ou ninhos (ALAM et al., 2010; SEIPKE; KALTENPOTH; HUTCHINGS, 2012). Adicionalmente, isolamentos de *Streptomyces* no intestino de térmitas, minhocas e outros invertebrados tem demonstrado a produção de antibióticos e enzimas que ajudariam no processo de degradação de matéria orgânica, assim como defesa contra patógenos, mas até o momento não se estabeleceu a profundidade da relação simbiótica ou o papel da bactéria *in-vivo* (SEIPKE; KALTENPOTH; HUTCHINGS, 2012).

2.1.3 Características do genoma de *Streptomyces* spp.

Diferente da maioria dos genomas bacterianos, o cromossomo de *Streptomyces* forma estruturas lineares que adicionalmente, contém sequências repetidas invertidas nos extremos (Terminal Inverted Repeats - TIRs) carregando proteínas terminais covalentes (Tp), formando uma estrutura telomérica característica em genomas virais e plasmídeos lineares (HOPWOOD, 2006; LEZHAVA et al., 1995; NINDITA et al., 2015). Estas proteínas são codificadas pelo operon *tpg-tap* ou homólogos deste, que se encontra dentro destes extremos terminais invertidos.

Por outra lado, o cromossomo das bactérias do gênero *Streptomyces* abrange ao redor de 8 Mb (BENTLEY et al., 2002; IKEDA et al., 2003), sendo o menor genoma reportado de 6 Mb (ZABURANNYI et al., 2014), e o maior de 11 Mb (WANG et al., 2013). Este tamanho é surpreendentemente grande, comparado com outros genomas bacterianos como a *Escherichia coli* (4.6 Mb) ou *Staphylococcus aureus* (2.9 Mb). Outra característica destes genomas é o alto conteúdo de guanina e citosina, ao redor de 70%, modificando a preferência de códons iniciadores nos genes. Classicamente o códon ATG que codifica para Metionina é o dominante nas ORFs preditas em genomas bacterianos, mas devido ao alto conteúdo G+C% em *Streptomyces*, um maior número de genes utiliza o códon GTG que codifica para N-formilmetionina, apresentando uma variável a mais para a predição de ORFs empregando ferramentas bioinformáticas (HOPWOOD, 2006).

Dentro deste grande cromossomo linear, a região central geralmente contém genes responsáveis pelo metabolismo primário, replicação do DNA e divisão celular, sendo uma região altamente conservada. Em contrapartida, os extremos do cromossomo, caracterizados pela alta instabilidade genética que facilita a inserção de elementos moveis, albergam um grande número de genes envolvidos no metabolismo secundário, incluindo clusters biossintéticos de produtos naturais (AIGLE et al., 2014; DYSON, 2011).

2.2 Metabólitos secundários

2.2.1 Aspectos gerais

Os metabólitos secundários são compostos de baixo peso molecular que geralmente exibem atividade biológica. Na natureza eles desempenham papéis diferentes como toxinas, reguladores, moléculas sinalizadoras, antibióticos, entre outros (DEMAIN; ADRIO, 2008).

A descoberta e uso destes metabólitos têm ajudado a dobrar a expectativa de vida dos humanos no século 20, combatendo dores, enfermidades infecciosas, revolucionando a prática da medicina. Os produtos naturais de origem microbiana incluem antibióticos contra bactérias e fungos, antitumorais, imunossuppressores, hipocolesterolêmicos, antiparasitários, herbicidas, reguladores do crescimento vegetal, pesticidas, inseticidas, antivirais; e atualmente está se estudando seu uso em tratamentos contra obesidade, enfermidades causadas por príons, diabetes, hiperlipoproteinemia, úlceras gástricas e hemocromatose em pacientes de diálise; entre outros (DEMAIN, 2014).

A diversidade de produtos naturais e os esqueletos químicos que as formam, tem sido uma inspiração para o desenvolvimento de fármacos inovando no desenho de mecanismos para modificar estas moléculas, tornando-as menos tóxicas e mais efetivas. Isto tem motivado a contínua busca de novos compostos explorando novos nichos ecológicos e diversos táxons aumentando a diversidade de produtos naturais descritos até o momento (WALSH; FISCHBACH, 2010). A biossíntese da maioria destes se encontra codificada no genoma dos microrganismos na forma de agrupações gênicas ou clusters biossintéticos, que contêm genes relativamente grandes, alguns com estrutura modular responsáveis pela síntese de diversas

famílias de produtos como os policetídeos (PKS), peptídeos não ribossomais (NRPS), aminociclitois, peptídeos de síntese ribossomal (RiPPs), siderofóros, terpenos, ectoínas, fosfonatos, entre outros.

2.2.2 Policetídeos PKS

Os compostos policetídeos destacam-se dentro da grande variedade de produtos naturais, devido à ampla gama que exibem em quanto a funcionalidade e estruturas (Figura 2). Além disso, estes produtos representam 1/3 dos fármacos aprovados entre o 2005-2007 e ao mesmo tempo, muitos continuam avançando nos ensaios clínicos. Uma evidência da sua importância é que as vendas de medicações derivadas de policetídeos alcançam os 20 bilhões de dólares (WEISSMAN, 2009).

Os policetídeos são compostos amplamente distribuídos na natureza produzidos por enzimas multimodulares chamadas de policetídeo sintases, e estes compartilham as mesmas etapas dentro do mecanismo da biossíntese de ácidos graxos, como as repetitivas elongações. Dependendo da estrutura e função, as PKS podem classificar-se em três tipos; as Tipo I são sintetizadas por várias proteínas multimodulares formando uma única unidade, alinhadas especificamente para a elongação da cadeia policetídica. As moléculas do Tipo II pertencem a um complexo de enzimas multimodulares que agem de forma iterativa para originar o centro da molécula (FISCHBACH; WALSH, 2006; WALSH; FISCHBACH, 2010). Estes dois tipos de compostos são amplamente encontrados em microrganismos, diferentemente das moléculas PKS do Tipo III, que são mais abundantes em plantas, mas têm sido descritos em alguns microrganismos. Estas últimas são sintetizadas por proteínas homodiméricas (HWANG et al., 2014). Dentre os policetídeos, encontram-se compostos de importância clínica como a eritromicina, tetraciclina, doxorubicina, daunorubicina, rapamicina, anfotericina B, griseofulvina, lovastatina e o macrolídeo FK506 (NEWMAN; CRAGG, 2012; WALSH; FISCHBACH, 2010).

Figura 2 Diversidade estrutural de moléculas policetídeas

Fonte: (DUNN; KHOSLA, 2013)

2.2.3 Peptídeos não ribossomais NRPS

Os antibióticos da classe dos peptídeos não ribossomais (NRPS) são produzidos por um conjunto de enzimas multifuncionais caracterizadas como sintases de peptídeos não ribossomal (NRPS). Estas enzimas contêm vários módulos encarregados de adenilar (domínio A: Adenilação), condensar (domínio C: Condensação) e sintetizar (domínio T: Tiolação, E: Epimerização) moléculas antibióticas utilizando diferentes aminoácidos como unidades de construção (Figura 3) ((WALSH; FISCHBACH, 2010; HWANG et al., 2014). Ao contrário dos peptídeos, que são sintetizados ribossomalmente, os NRPS possuem características estruturais únicas, como a presença de D-aminoácidos, cadeias de ácidos graxos ligadas ao N-terminal, aminoácidos N- e C-metilados, resíduos N-formilados, elementos heterocíclicos, glicosilados e resíduos fosforilados (FINKING; MARAHIEL, 2004).

Figura 3 Módulos enzimáticos para a síntese de peptídeos não ribossomais pela condensação de aminoácidos

Fonte: Modificado de (WALSH; FISCHBACH, 2010)

A diversidade das moléculas, devido às enzimas multimodulares, e a variedade de aminoácidos e derivados incorporados na síntese, fazem com que as atividades biológicas sejam igualmente diversas, como dano na membrana ou

parede celular, ou inibição da síntese de proteínas pelo dano ao DNA ou RNA, tendo uma aplicação variada como antibacterianos, antifúngicos, antivirais e antitumorais (FELNAGLE et al., 2008).

Exemplos deste tipo de compostos são os antibióticos β -lactâmicos, como a penicilina e a cefalosporina; os antibióticos lipopeptídicos cíclicos como a daptomicina usada contra infecções de Gram positivos multi-resistentes; antifúngicos como a ciclosporina e os glicopeptídeos como a vancomicina e a teicoplanina (FELNAGLE et al., 2008; SIEBER; MARAHIEL, 2003).

2.2.4 Peptídeos Ribossomais RiPPs

Estas moléculas são produzidas nos três domínios da vida, os genes biossintéticos são ubíquos nos genomas atualmente sequenciados e a diversidade estrutural é vasta. Como vantagem, os peptídeos sintetizados pelos ribossomos sofrem modificações pós-traducionais que diversificam a estrutura destes produtos naturais, permitindo melhor reconhecimento dos alvos, aumentando a estabilidade química e metabólica, e a sua vez melhorando a funcionalidade química (ARNISON et al., 2013).

Os peptídeos ribossomais derivam de curtos precursores, ao redor de 100 aminoácidos, que são enzimaticamente modificados pós e co-traducionalmente. Geralmente o peptídeo precursor contém uma sequência relativamente conservada, que é chamada de peptídeo ou sequência líder N-terminal, que facilita o reconhecimento dos transportadores e exportação da molécula. O extremo C-terminal do peptídeo precursor codifica a sequência principal que é modificada enzimaticamente (*core sequence*) (SCHMIDT, 2010). Usualmente o peptídeo líder é clivado da sequência modificada, gerando um produto peptídico curto chamado de peptídeo maduro (Figura 4).

Figura 4 Via biossintética geral para os peptídeos ribossomais (RiPPs).

Fonte: Modificado de (ARNISON et al., 2013)

2.2.5 Terpenos

Ao redor de 50.000 metabólitos terpenoides têm sido isolados de plantas terrestres, plantas marinhas e fungos, mas poucos têm sido achados em procariontes. As primeiras moléculas encontradas em bactérias foram os terpenos odoríferos, principalmente, a geosmina, um sesquiterpeno degradado, que proporciona o odor característico de terra húmida (CANE; IKEDA, 2012). Os terpenos podem ser divididos em monoterpenos cíclicos, sesquiterpenos e diterpenos. Estes são formados por variações do mecanismo de ciclização universal, iniciando pela ionização enzimática dos precursores acíclicos geranil difosfato (GPP), farnesil difosfato (FPP) e geranil-geranil difosfato (GGPP) (Tabela 1) (YAMADA et al., 2015).

Tabela 1 Metabólitos terpenóides amplamente estudados.

Precursor	Classe de terpeno	Moléculas
GPP	Monoterpenos	Limoneno – Geraniol
FPP	Sesquiterpenos e Triterpenoides	Geosmina – Esqualeno
GGPP	Diterpenos e Politerpenos	Terapentecina – Giberelinas

As terpeno sintases, são as enzimas encarregadas de converter os precursores GPP, FPP e GGPP em compostos cíclicos dando origem à grande variedade estrutural (Figura 5). Tradicionalmente, os genes que codificam estas enzimas são altamente conservados no nível de aminoácidos em plantas e fungos; no entanto, em bactérias, estas enzimas apresentam um grande desafio para as ferramentas bioinformáticas que estão disponíveis atualmente, já que não só as sequências são diferentes às das plantas e fungos, mas possuem relativamente baixa similaridade mútua entre as sequências (YAMADA et al., 2015).

Figura 5 Estruturas dos terpenos mais produzidos por bactérias

Fonte: (YAMADA et al., 2015)

Os terpenos têm um papel importante na ecologia química das plantas, insetos, fungos e bactérias. Eles atuam como sinalizadores, ou mecanismos de defesa indiretos contra patógenos, e atualmente os terpenos de origem microbiana têm sido usados como potentes agentes antimicrobianos como o pentaleneno ou a terapentecina (CANE; IKEDA, 2012b; YAMADA et al., 2015). Adicionalmente, existem estudos de modificações à via metabólica de biossíntese de terpenos em *Streptomyces venezuelae*, induzindo a produção de bisaboleno, um terpeno que tem sido analisado como possível bio-combustível (PHELAN et al., 2014).

2.2.6 Aminociclitolis

Os aminociclitolis são um grupo de moléculas de origem microbiano, isolados geralmente de actinobactérias de grande importância clínica e agrícola. Estes podem ser divididos, dependendo da sua estrutura, em três grandes grupos (Tabela 2). Os primeiros são os antibióticos aminoglicosídeos e os aminociclopentinóis os quais, atuam interagindo com o RNA ribossomal de bactérias impedindo a síntese de proteínas, ação devida a diferentes deoxi e amino açúcares ligados ao centro de aminociclitol da molécula antibiótica (KUDO; EGUCHI, 2009; KUDO et al., 2007; MAHMUD, 2009). O último grupo está representado pelos C₇-N aminociclitolis que atuam mimetizando a estrutura da glicose ou outros açúcares, usando o anel de aminociclitol inibindo a ação da β -glicosidase ou enzimas relacionadas como α -amilases (SCHWIENTEK et al., 2012).

Tabela 2 Produtos naturais da família dos aminociclitolis

Tipo de Aminociclitol	Produtos
Aminoglicosídeos	Gentamicina, Canamicina, Estreptomicina
Aminociclopentinóis	Pactamicina, Allosamidina
C₇-N Aminociclitolis	Acarbose, Validamicina, Salbostatina

Biossinteticamente, estas moléculas derivam de unidades simples de açúcares, formadas geralmente por um grupo de enzimas conhecidas como açúcar-fosfato ciclasas. Estas incluem a 1L-mio-inositol 1-fosfato sintase (MIPS), a 2-desoxi-cilo-inosose sintase (DOIS) e a 2-epi-5-epi-valiolone sintase (EVS), e a função delas é catalisar a ciclização dos açúcares fosfatados para produzir uma variedade de intermediários de ciclitol (Figura 6).

Figura 6 Via biossintética geral para as moléculas derivadas do aminociclitol.

2.2.7 Fosfonatos

O fósforo é um elemento essencial para todos os organismos presente na maioria de compostos básicos da célula (DNA, ATP, fosfolipídios) nos estados mais oxidados na forma de fosfato inorgânico, ésteres de fosfato e fosfoanidridos. Uma exceção é a formação de compostos com enlaces C-P chamados fosfonatos. Estas moléculas são quase isotéricas aos ésteres fosfato, no entanto, possuem uma maior estabilidade e são resistentes a tratamentos químicos como ebulição e presença de ácidos ou bases fortes. Elas têm sido subexploradas mesmo possuindo uma alta atividade biológica antibiótica, antifúngica e antiparasitária (Figura 7). Hoje em dia, algumas destas moléculas são amplamente usadas como herbicidas (Glifosato) e em tratamentos médicos contra a osteoporose (Alendronato) (YU et al., 2013).

Figura 7 Diversidade estrutural e biológica dos fosfonatos produzidos por bactérias.

Fonte: Modificado de (METCALF; VAN DER DONK, 2009).

O primeiro produto natural da família dos fosfonatos identificada foi o 2-aminoetilfosfonato (AEP) encontrado em um grande número de organismos como análogos aos lipídeos de membrana, sugerindo como função biológica a estabilidade e fluidez das membranas (METCALF; VAN DER DONK, 2009).

A biossíntese do AEP (Figura 8) começa do fosfoenolpiruvato (PEP) como precursor sendo transformado a fosfonopiruvato (PnPy) pela enzima PepM (Fosfoenolpiruvato fosfomutase), seguido pela conversão a fosfonoacetaldeído (PnAA) por uma enzima descarboxilase e finalizando pela transaminação envolvendo a enzima AEP-transaminase (aminotransferase) (YU et al., 2013).

Figura 8 Síntese do Aminoetilfosfonato, precursor dos compostos C-P.

Fonte: Modificado de (METCALF; VAN DER DONK, 2009).

A presença do gene *pepM*, que codifica para a fosfoenolpiruvato fosfomutase, tem sido amplamente utilizado como *screening* para a identificação *in silico* de produtos fosfonatados, facilitando a descoberta destes produtos naturais. Além disso, já que os genes que codificam para a produção de fosfonatos geralmente encontram-se agrupados, as análises de vizinhança gênica do gene *pepM*, sugerem que a diversidade estrutural e funcional destes produtos é maior do que reportado atualmente (JU; DOROGHAZI; METCALF, 2014; YU et al., 2013).

2.3 *Streptomyces* como fábrica de produtos naturais

O phylum *Actinobacteria* é altamente conhecido por ser uma fonte prolífica de produtos naturais, especialmente o gênero *Streptomyces* sp. Ao invés dos genes de metabolismo primário que estão distribuídos pelo genoma todo, os genes de metabolismo secundário agrupam-se em clusters gênicos, que no caso dos *Streptomyces*, localizam-se nos extremos do cromossomo. Estes clusters são compostos de genes encarregados da biossíntese da molécula, genes para produção de enzimas pós-modificadoras, genes reguladores e genes transportadores ou de resistência.

Um dos primeiros *clusters* biossintéticos a ser descrito em *Streptomyces* foi o antibiótico actinorodina presente em *S. coelicolor*, e desde o sequenciamento do genoma tornou-se claro que o potencial biossintético das actinobactérias estava sendo subestimado (BENTLEY et al., 2002; COLE et al., 1987). Atualmente se sabe que no cromossomo dos microrganismos existem ao redor de 20 a 30 clusters biossintéticos, que incluem produtos utilizados na clínica como a daunorubicina, codificada no genoma de *S. peucetius* (VASANTHAKUMAR; KATTUSAMY; PRASAD, 2013); a rapamicina em *S. rapamycinicus* (BARANASIC et al., 2013; SCHWECKE et al., 1995); e o agente imunossupressor FK506, pertencente a *S. tsukubaensis* (CHEN et al., 2012; MOTAMED; SHAFIEE, 1998).

Estes novos *clusters* são encontrados na sua maioria como vias crípticas ou silenciadas em condições de cultura padrão em laboratório (NETT; IKEDA; MOORE, 2009; OCHI; HOSAKA, 2013). O desenvolvimento de metodologias para a expressão destes compostos tem levado à interdisciplinaridade, integrando ferramentas ômicas, para a elucidação da rota metabólica e a estrutura química das

possíveis moléculas (AIGLE et al., 2014; BALTZ, 2010; GOMEZ-ESCRIBANO; BIBB, 2014).

2.4 Regulação de vias biossintéticas de metabólitos secundários e ativação de vias crípticas

A regulação do metabolismo de *Streptomyces* tem sido amplamente estudada devido à importância industrial do gênero. A diferenciação morfológica e a produção de antibióticos são ativadas quando há uma falta de nutrientes ou mudanças ambientais, diminuindo o rápido crescimento vegetativo. Sob estas circunstâncias uma série de mudanças na expressão global ocorre, sendo isto mediado por mecanismos regulatórios complexos.

As primeiras análises realizadas aos clusters biossintéticos demonstraram que usualmente estas vias possuem genes reguladores, que geralmente têm um efeito maior nos níveis de produção desse metabólito secundário. Este fato dividiu as moléculas reguladoras presentes em *Streptomyces* em dois grupos; o primeiro grupo são as chamadas globais ou reguladores pleiotrópicos, que controlam diversas vias metabólicas e não se encontram relacionados especificamente a um tipo de *cluster* biossintético. Estes genes reguladores podem ter sido afetados por fatores ambientais como pH, temperatura ou nutricionais, como a limitação de nitrogênio, fósforo ou a presença de diversos compostos no meio (BIBB, 2005; MARTÍN; LIRAS, 2010). Um exemplo desta regulação pleiotrópica é a presença do gene *adpA* em *Streptomyces* que codifica para a proteína Bld-H que intervém na produção de micélio aéreo, na esporulação e no controle da replicação do cromossomo. Além disso, atua como ativador ou repressor indireto para a produção de compostos como estreptomicina, ácido clavulânico, nikkomicina, dentre outros (GUYET et al., 2013; LIU et al., 2013; VAN WEZEL; MCDOWALL, 2011; WOLAŃSKI; JAKIMOWICZ; ZAKRZEWSKA-CZERWIŃSKA, 2012). O entendimento destes mecanismos regulatórios globais e sua interação com os nutrientes e fatores abióticos, têm auxiliado no desenvolvimento de condições apropriadas para a indução da expressão de *clusters* crípticos (RABYK; OSTASH; FEDORENKO, 2014).

O segundo grupo de reguladores são os reguladores específicos de vias metabólicas que, geralmente, estão localizados dentro de *clusters* de metabólitos

secundários junto a genes responsáveis pela biossíntese, modificações posteriores, genes de resistência e transporte (AIGLE; CORRE, 2012; BIBB, 2005; MARTÍN; LIRAS, 2010).

Estes reguladores estão divididos em diferentes famílias, sendo as maiores, a família LAL (large ATP-binding regulators of the Lux-R Family) e a família SARP (*Streptomyces* Antibiotic Regulatory Proteins). Esta última família mencionada, possui um domínio de ligação ao DNA OmpR-like adicional a um domínio ativador transcricional do grupo de bactérias (BTAD). São amplamente conhecidos, já que fazem parte do final de uma cadeia reguladora e são responsáveis pela ativação de *clusters* biossintéticos, como diversos tipos de policetídeos e outros metabólitos secundários. Além disso, estes genes têm sido encontrados exclusivamente em bactérias filamentosas. Os reguladores tipo SARP reconhecem sequências consenso, localizadas a oito nucleotídeos antes da região -10 do promotor (AIGLE; CORRE, 2012; MARTÍN; LIRAS, 2010). Um exemplo deste tipo de ativador é o gene *actII-ORF4* de *Streptomyces coelicolor* que se encontra dentro do *cluster* para produção de actinorodina, um antibiótico tipo PKS. Este gene regula o momento em que a via deve ativar-se, a síntese do composto e as quantidades do policetídeo, afetando, assim, a transcrição e tradução dos genes. A sua região promotora é alvo para a ligação de outras moléculas reguladoras que atuam não somente na via, mas, também, na regulação global (LIU et al., 2013; MARTÍN; LIRAS, 2010). O mecanismo de ação deste gene já tem sido descrito por sua complexidade na regulação específica dos *clusters* biossintéticos, ajudando, assim, numa melhor caracterização da família de proteínas reguladoras tipo SARP.

2.5 Potencial biotecnológico de *Streptomyces olindensis* DAUFPE 5622

Na década de 60, foi isolada a bactéria *S. olindensis* a partir de solo da cidade de Olinda, pelo Departamento de Antibióticos da Universidade Federal de Pernambuco (DAUFPE) (BIEBER et al., 1989) tornou-se interessante pela detecção da produção de um composto antitumoral policetídeo de cor roxa, com a presença de três açúcares ligados ao carbono 10, nomeada retamicina (Figura 9) (BIEBER et al., 1989). Estudos posteriores demonstraram que a molécula possuía seis açúcares ao invés de três (Figura 6), convertendo-se numa molécula de grande interesse pelo interessante padrão de glicosilação (FURLAN et al., 2004; GARRIDO et al., 2006)

Figura 9 (A) Estrutura da Retamicina detectada por Bieber et al (1989), (B) Estrutura da Cosmomicina D detectada por Furlan et al (2004)

Fonte: Modificado de (BIEBER et al., 1989; FURLAN et al., 2004)

2.5.1 Biossíntese da Cosmomicina D em *Streptomyces olindensis* DAUFPE 5622

A construção de uma biblioteca de cosmídeos realizada por Garrido (2006) atribuiu a síntese parcial da Cosmomicina D a uma agrupação gênica de aproximadamente 14 kb, que possuía 13 ORFs completas e 1 ORF incompleta que cumpriam as funções de biossíntese do policetídeo, açúcares, glicosiltransferases, genes reguladores e genes de transporte da molécula. Posteriormente o sequenciamento do genoma foi realizado (CONTRERAS, 2013; ROJAS et al., 2014), e identificou-se um fragmento de 45 kb responsável pela síntese total da molécula (Figura 10).

Figura 10 Agrupação gênica para a biossíntese da Cosmomicina D em *Streptomyces olindensis* DAUFPE 5622

Fonte: (CONTRERAS, 2013).

A partir dos trabalhos de Garrido et al. (2006) e Contreras (2013) a via biossintética da Cosmomicina D foi elucidada (Figura 11), atribuindo função as ORFs detectadas *in-silico*, com a exceção das ORFs localizadas aos extremos do cluster e dois ORFs internas (ORF44 e *cosY*).

A biossíntese da molécula começa com a condensação de unidades iniciadoras de Propionil-CoA e Malnoil-CoA pelo PKS mínimo (*cosA*, *cosB*, *cosC*, *cosE* e *cosF*) (Figura 11). Partindo disto, uma serie de reduções e ciclizações são realizadas (*orf40*, *orf41*, *orf37*, *cosX*); e posteriormente o ácido aplanônico será modificado com a ação de metiltransferases, hidroxilases e reductases (*orf39*, *orf46*, *orf48*, *orf49*), que permitirá a entrada das unidades glicosiladas à molécula, conferindo a alta atividade biológica da cosmomicina D, portadora de um padrão pouco comum de glicosilação.

Adicionalmente ORFs envolvidas no transporte e resistência à molécula tem sido caracterizadas, elucidando o mecanismo da bactéria para reparar os danos que o composto possa causar e a expulsa (CONTRERAS, 2013). Outras ORFs, tais como a *cosS* (ORF 52), pensa-se que atua como reguladora da expressão da cosmomicina D, atuando como regulador específico da via.

Figura 11 Via biossintética da cosmomicina D em *Streptomyces olindensis* DAUFPE 5622

3 MATERIAIS E MÉTODOS

3.1 Análise genômica *in-silico*

3.1.1 Anotação e análise do genoma de *Streptomyces olindensis*

A sequência do genoma foi obtida previamente utilizando a tecnologia de piro-sequenciamento 454 usando o software Roche GS FLX (v 2.6), gerando 120 *Contigs* com o total de 9.4 Mb com um N50 de 246kb (CONTRERAS, 2013; ROJAS et al., 2014). Os *reads* originados foram montados usando Mira e Newbler (v 3.2/2.6) e as ORFs foram preditas com o software preditor Glimmer3 (DELCHER et al., 2007). Este software foi treinado criando um modelo de contexto com interpolação (ICM - Interpolated Context Model), utilizando sequências de diversos genes dentro do gênero *Streptomyces* que tinham como códon de iniciação não só ATG, mas também GTG e TTG, presentes frequentemente nos genes de actinobactérias devido ao alto conteúdo de G+C% no genoma destas (HUTCHINGS et al., 2006).

Finalmente, estas predições foram anotadas automaticamente seguindo o *pipeline* do servidor RAST (AZIZ et al., 2008), que usa a base de dados do SEED Project, para comparar os genes anotados e construir subsistemas para a classificação das proteínas preditas (OVERBEEK et al., 2014).

Posteriormente, o genoma foi reanotado no *pipeline* EGene (DURHAM et al., 2005), desenvolvido pelo grupo de pesquisadores do Departamento de Parasitologia no Instituto de Ciências Biomédicas e do Departamento de Ciências da Computação do Instituto de Matemática e Estatística da Universidade de São Paulo, equipe liderada pelo Dr. Arthur Gruber e Dr. Alan M. Durham. Este *pipeline* utiliza BlastP (JOHNSON et al., 2008) para atribuir funções aos potenciais genes e InterPro (MITCHELL et al., 2015) para detectar domínios específicos nas proteínas preditas. Adicionalmente, utilizou-se o software AlienHunter (VERNIKOS; PARKHILL, 2006) para prever possível transferência horizontal no genoma e, para identificar proteínas com possíveis domínios transmembranares e sítios de clivagem, foram usados Phobius (KÄLL; KROGH; SONNHAMMER, 2004), TMHMM (KROGH et al., 2001) e SignalP (PETERSEN et al., 2011). Ao mesmo tempo, cada gene foi assignado dentro de um grupo de ortólogos utilizando a base de dados eggNOG

(POWELL et al., 2014) para identificar as funções metabólicas destes genes e a proporção destes dentro do genoma.

3.1.2 Análise filogenética de *Streptomyces olindensis* DAUFPE 5622

Foram realizadas duas análises filogenéticas, a primeira usando o gene 16S rDNA pertencente a *S. olindensis* e genes de rDNA 16S actinobactérias já depositadas no NCBI. O alinhamento foi feito com o algoritmo do software MUSCLE (EDGAR, 2004) no programa MEGA6 (TAMURA et al., 2013). A árvore filogenética foi construída usando o método de Máxima Verossimilhança com um bootstrap de 1000 réplicas, e as distâncias evolutivas foram obtidas pelo modelo Kimura 2-parâmetros. Como grupos externos utilizaram-se as linhagens *Thermobifida fusca* e *Micromonospora peucetia* DSM 43363.

A segunda árvore que fornece contexto filogenético para a análise de metabolismo secundário, foi construída utilizando o gene rDNA 16S pertencente a *S. olindensis* concatenado com quatro genes clássicos conservados (*rpoB*, *gyrB*, *atpD*, *recA*); alinhando-os com genes de espécies relevantes de *Streptomyces* produtoras de antraciclinas. Como grupo externo utilizou-se a linhagem *Micromonospora aurantiaca* ATCC 27029. O alinhamento foi feito com o algoritmo do software MUSCLE (EDGAR, 2004) no programa MEGA6 (TAMURA et al., 2013). A árvore filogenética foi construída usando o método de Máxima Verossimilhança com um Bootstrap de 1000 réplicas, e as distâncias evolutivas foram obtidas pelo modelo Kimura 2-parâmetros.

3.1.3 Genômica comparativa entre espécies de *Streptomyces*

Para comparar a diversidade das vias metabólicas entre linhagens de *Streptomyces* realizou-se um gráfico com o número de clusters biossintéticos de cada espécie representada na árvore filogenética concatenada, que os autores reportaram ou preditos pelo software antiSMASH (WEBER et al., 2015), e observou-se a variedade de moléculas que potencialmente estes micro-organismos podem produzir.

Adicionalmente, foi comparado o genoma de *Streptomyces olindensis* com outras espécies, baseando nas categorias funcionais da base de dados de ortólogos

do KEGG (KO) (MAO et al., 2005). As regiões codificadoras foram assignadas a estas categorias e foram agrupadas conforme a sua função metabólica. Realizou-se um diagrama de Venn para identificar o número de categorias que os cinco genomas compartilham, formando o *core* genoma, e categorias funcionais exclusivas de *Streptomyces olindensis* visando identificar o potencial biotecnológico desta linhagem.

3.1.4 Identificação e anotação de clusters biossintéticos envolvidos na produção de metabólitos secundários

Para a detecção dos *clusters* biossintéticos de metabólitos secundários foi utilizado o servidor web AntiSMASH (WEBER et al., 2015) que permitiu não só a detecção de domínios conservados para a determinação da presença de agrupações gênicas, mas também fez uma busca de homologia de *clusters* em outros organismos baseando-se no algoritmo do *ClusterBlast*. Baseados nisto, as sequências foram curadas manualmente empregando o software Delta-BLAST (JOHNSON et al., 2008) e comparando as arquiteturas dos domínios com o HMMR (EDDY, 2011; FINN; CLEMENTS; EDDY, 2011).

3.2 Super-expressão dos genes ativadores SARP nas vias biossintéticas de produtos policetídeos e do aminociclitol.

3.2.1 Desenho dos iniciadores

Um conjunto de iniciadores foi desenhado para amplificar os genes reguladores da família SARP dos *clusters* biossintéticos, escolhidos para sua superexpressão e posterior produção em *S. olindensis*, no software Primer3 (UNTERGASSER et al., 2012). Os iniciadores contêm a sequência das enzimas de restrição a usar (*Xba*I e *Eco*RI), para a clonagem no vetor (APÊNDICE A).

3.2.2 Crescimento das linhagens utilizadas

Para a identificação das vias biossintéticas foi usada a linhagem de *Streptomyces olindensis* DAUFPE 5622 mantida em condições padrão de crescimento (28 °C em agitador orbital a 180 rpm por um período de 3 a 5 dias; tanto

para crescimento em meio TSB, como para produção de cosmomicina D em meio R5M). Para a extração de DNA total, a linhagem foi incubada por 36 horas em meio YEME líquido. Por outra parte, o cultivo padrão de *Escherichia coli* foi realizado em meio líquido LB a 37 °C em estufa por 16-20 horas a 150 rpm.

3.2.3 Conservação das linhagens

A linhagem de *Streptomyces olindensis* foi preservada através da suspensão de esporos em glicerol 20%, recolhidos de cultivos crescidos durante 4-6 dias em meio A, meio aveia e R5M e armazenados a -20 °C. A preservação das linhagens de *E. coli* se realizou partindo de 1 ml de uma cultura crescida *overnight*, após centrifugar a 2900 rpm por 10 minutos o sedimento foi ressuspensionado em meio LB fresco com glicerol 20% e, posteriormente, armazenadas a -80 °C e -20 °C.

3.2.4 Extração de DNA total

Para a extração de DNA *Streptomyces olindensis* genômica e plasmidial foram utilizados os protocolos descritos por Kieser et al. (2002). A obtenção do DNA plasmidial de *E. coli* em pequena escala foi realizada pelo método de lise alcalina descrito por Sambrook et al. (1989).

3.2.5 Reações de PCR

As reações de PCR foram realizadas utilizando a enzima *Pfu* DNA polimerase da Thermo Scientific®, adicionando DMSO 10% para aumentar a T_m dos primers, ideal para *amplicons* com alto conteúdo de G+C%. Usou-se um termociclador Mastercycler 5331 Gradient (Eppendorf®, Hamburgo, Alemanha), e as corridas foram feitas sob as seguintes condições: um ciclo inicial de 95 °C por 3 minutos; 30 ciclos de 20 segundos a 95 °C; 20 segundos a 60 °C; 2 minutos a 72 °C mais 4 segundos por ciclo, seguido de uma extensão final de 8 min. O DNA amplificado foi purificado empregando o KIT GFX MicroSpin Column (General Electric, Buckinghamshire, UK) seguindo as recomendações indicadas pelo fabricante.

3.2.6 Transformação de *E. coli* e protoplastos de *S. olindensis*

A ligação dos fragmentos foi feita utilizando o vetor de expressão constitutiva PEM4A (9.40 kb), escolhido devido a sua estabilidade dentro de *E. coli* e *Streptomyces* sp., e também a presença do promotor constitutivo forte ermE*. O plasmídeo foi digerido utilizando as enzimas XbaI e EcoRI (Thermo Scientific ©, Waltham, Massachusetts, Estados Unidos) de acordo com o protocolo do fabricante. Após a digestão foi feita a ligação utilizando T4 DNA Ligase (New England Biolabs® Inc. Ipswich, Massachusetts, Estados Unidos). A transformação de *E. coli* DH10β foi realizada utilizando eletroporação seguindo o protocolo descrito por Sambrook et al.(1989). Após a seleção foi realizada a extração do plasmídeo com o inserto utilizando o protocolo descrito no item 4.2.4. Posteriormente, devido ao mecanismo de resistência a metilação restrito do *S. olindensis*, o plasmídeo PEM4A foi clonado na *E. coli* ET12567/pUB307 para de-metilar e evitar a rejeição na transformação de protoplastos seguindo o protocolo descrito por Sambrook et al.(1989). Os transformantes foram selecionados após 24 horas e o DNA plasmidial foi extraído como foi descrito no item 4.2.4. Para superexpressar os genes reguladores foram utilizados protoplastos da linhagem *Streptomyces olindensis* DAUFPE 5622, utilizando o protocolos descrito por Kieser et al. (2002).

4 RESULTADOS E DISCUSSÃO

Streptomyces olindensis DAUFPE 5622 foi isolada nos anos 60 a partir de uma amostra de solo e, desde então, tem sido amplamente estudada devido à produção de uma antraciclina do grupo dos polipeptídios denominada cosmimicina D, possuindo uma alta atividade antitumoral, que é de grande interesse pelo complexo padrão de glicosilação pouco comum (GARRIDO et al., 2006). Alguns genes do *cluster* para a produção deste antitumoral foram encontrados por Garrido (2006), mas para fechar o conjunto gênico decidiu-se sequenciar o genoma da bactéria, intuindo-se um grande potencial gênico e metabólico.

4.1 Generalidades físicas do genoma de *Streptomyces olindensis* DAUFPE 5622

Para a descrição das características físicas do genoma de *Streptomyces olindensis* empregaram-se diferentes pipelines para a predição e anotação de *Open Reading Frames* (ORFs). Inicialmente o software Glimmer3 foi utilizado para a predição de ORFs que emprega Modelos Interpolados de Markov para distinguir entre as sequências com regiões codificadoras daquelas que não têm; determinando assim, o início e fim das ORFs, o sentido de transcrição e os códons de iniciação (DELCHER et al., 2007). Previamente foi usado um conjunto de treinamento que possuía diversos genes já descritos de bactérias do gênero *Streptomyces*, para a criação do modelo de contexto com interpolação (ICM) para melhorar a detecção de genes em *Streptomyces*. Este passo foi essencial por dois motivos: primeiro, para a determinação dos códons nas ORFs devido ao alto conteúdo G+C que varia entre 69 e 74%, influenciando na utilização preferencial de códons com alto conteúdo de G-C, especialmente na terceira posição (WRIGHT; BIBB, 1992). Segundo, a detecção de ORFs curtas comumente encontradas em actinobactérias como os genes codificadores para proteínas carreadoras de acilo (ACP) ou as proteínas carreadoras de peptidil ou fosfopanteteína (PP), essenciais na síntese de metabólitos secundários de origem policetílica ou não ribossomal (WANG et al., 2014).

Após a predição de ORFs, foi empregado o servidor RAST para a anotação automática funcional do genoma, *pipeline* amplamente usada na anotação de genomas procariotas classificando as regiões codificadoras, em subsistemas

funcionais, contendo diversas famílias de proteínas dentro destes (APÊNDICE B) (OVERBEEK et al., 2014). Este tipo de plataforma representa uma solução aos grandes volumes de dados gerados e propõe uma anotação funcional e interativa, que é ideal para a descrição de genomas bacterianos que são explorados para a obtenção de novos produtos metabólicos.

Posteriormente foi disponibilizado o *pipeline* EGene (DURHAM et al., 2005), que permitiu realizar análises globais do genoma, oferecendo a informação em formatos únicos e simples para facilitar as análises de genômica comparativa e anotação das vias de metabólitos secundários de interesse biotecnológico.

As características principais do genoma estão representadas na Tabela 3. O genoma de *S. olindensis* possui um plasmídeo de 76 kb localizado no contig 27 (Numero de acesso JJOH01000027), confirmado usando Eletroforese de Campo Pulsado (PFGE) (Dados não mostrados).

Tabela 3 Características gerais do genoma de *S. olindensis* DAUFPE 5622 comparado com outras espécies de *Streptomyces*

Espécies	Tamanho (pb)	G+C %	No. CDS	Tm médio CDS (pb)	No. rRNA	No. tRNA	No. Plasmídeos	Referencia
<i>S. olindensis</i> DAUFPE 5622	9.381.854	71.6	8.715	941	6	65	1	Presente estudo
<i>S. coelicolor</i> A3(2)	8.667.507	72.1	7.825	991	6	63	1	Bentley et al. 2002
<i>S. avermitilis</i>	9.025.608	70.7	7.574	1027	6	68	1	Ikeda et al. 2003
<i>S. albus</i> J1074	6.841.649	73.3	5.832	1011	7	66	-	Zaburanny et al. 2014
<i>S. bingchengensis</i> BCW-1	11.936.683	70.7	10.023	1030	6	66	-	Wang et al. 2013

O genoma possui 8.715 possíveis regiões codificadoras (CDS) proporcional ao tamanho de 9.381.854 pb e o tamanho médio destes CDS é de 941 pb, o que pode variar dependendo dos preditores automáticos de ORFs. Destas 8.715 CDS 68.3% tem uma função putativa atribuída e 31.7% tem uma função desconhecida. Adicionalmente o genoma de *S. olindensis* possui 6 operons de RNA ribossomal (5S, 16S, 23S), como a maioria de genomas de *Streptomyces*, e 65 RNAs de transferência representando putativamente os 20 aminoácidos.

Foi realizada uma comparação das características físicas utilizando os genomas de *S. coelicolor*, *S. albus*, *S. avermitilis* e *S. bingchenggensis*, encontrando uma relação proporcional do tamanho do genoma com o número de CDS preditas nos cinco genomas (Tabela 3). Além disso, foi evidenciado que características como o alto conteúdo de G-C nos genomas é conservado em bactérias do gênero (WRIGHT; BIBB, 1992).

Por outra parte, a linearidade do cromossomo e a instabilidade genética presente nos extremos deste cromossomo são propriedades interessantes e particulares dos *Streptomyces*. A topologia do cromossomo na maioria dos genomas de *Streptomyces* tem sido relacionada com a presença ou ausência dos genes *tpg* (*terminal protein gene*), *tap* (*telomere-associated protein gene*) e *ttrA* (*terminal transfer protein gene*), ou homólogos dentro das sequências terminais invertidas (TIRs).

Estudos realizados nas espécies *S. coelicolor*, *S. avermitilis*, *S. lividans* e *S. griseus* demonstram experimentalmente a topologia utilizando eletroforese de campo pulsado, isolamento dos fragmentos terminais do DNA e deleção das proteínas terminais (BENTLEY et al., 2002; HUANG et al., 1998; OMURA et al., 2001; YANG et al., 2013). Estas pesquisas têm associado à presença dos genes *tpg*, *tpa* e *ttrA* ou homólogos destes genes com a linearidade do cromossomo, e o sequenciamento tem facilitado a detecção *in-silico* destes. Estes genes estão envolvidos na formação dos telômeros arquetípicos em *Streptomyces*, e podem-se encontrar cópias tanto no cromossomo como nos plasmídeos.

Em *S. olindensis* foram identificadas duas cópias das ORFs homólogas aos genes *tpg-tpa*, localizados contiguamente, em dois *contigs* diferentes dentro do cromossomo, além, de um gene homólogo ao *ttrA*, (Tabela 4) sugerindo a linearidade deste. A presença destas ORFs concorda com o observado em *S. coelicolor* e *S. griseus* (BENTLEY et al., 2002; HUANG et al., 1998). Adicionalmente, foram encontradas cópias dos genes *tpg-tpa* no plasmídeo putativo de *S. olindensis* o que coincide com o encontrado em *S. rochei* (NINDITA et al., 2015), bactéria que tem um cromossomo linear e dois plasmídeos com as mesmas características, mas só possui genes homólogos aos *tpg-tpa* nos plasmídeos, indicando que estas proteínas são responsáveis pela linearidade do cromossomo assim, como dos plasmídeos.

Tabela 4 Genes associados à replicação dos telômeros no genoma de *S. olindensis* DAUFPE 5622

Gene	Localização	Locus Tag	%Identidade	Micro-organismo
tpg	Contig11 JJOH1000011.1	DF19_08255	74.46%	<i>S. ambofaciens</i> ATCC 23877 SAML0477
	Contig100 JJOH1000100.1	DF19_07200	73.42%	
tap	Contig11 JJOH1000011.1	DF19_08260	75.66%	<i>S. ambofaciens</i> ATCC 23877 SAML0478
	Contig100 JJOH1000100.1	ORF não anotada 4178-3255	66.96%	
ttrA	Contig39 JJOH01000039.1	DF19_33015	59.95%	<i>S. ambofaciens</i> ATCC 23877 SAMT0022

Outra das características físicas relevante nas actinobactérias é a diversidade gênica, e um dos fatores associados a este fenômeno, é a presença de genes que permitem a entrada de sequências genômicas que podem apresentar vantagens adaptativas. Na tabela 5 descrevem-se 22 genes putativos no genoma de *S. olindensis* que codificam para integrases e transposases, encontrados ao longo do genoma de forma intacta ou truncados. Estas proteínas estão associadas à plasticidade e mobilidade de elementos gênicos no genoma. A presença destes genes somada às sequências invertidas terminais do cromossomo, são um dos mecanismos responsáveis pela instabilidade genética. Este fenômeno, característico nas bactérias deste gênero, afeta as propriedades fenotípicas, geralmente de forma pleiotrópica, incluindo a diferenciação morfológica e a produção de metabólitos secundários (BLAŽIČ et al., 2012; VOLFF1; ALTENBUCHNER, 1998).

Tabela 5 Genes associados à síntese de integrases e transposases

Função putativa	Número de acesso
Transposases	WP_037758872, WP_037758699, WP_037759152, WP_031114607, WP_031117236, WP_037765314, WP_037765487, WP_051649171, WP_037767566, WP_037767720, WP_037768110, WP_037769817
	WP_037760166, WP_037758626, WP_037760300, WP_037760923, WP_037761138, WP_037765190, WP_051649048, WP_031117466, WP_037765585, WP_037768020
Integrases	

Analisando a vizinhança gênica destas ORFs, encontraram-se genes envolvidos na síntese de proteínas transportadoras ou bombas de efluxo, genes envolvidos na instabilidade genética e adicionalmente, o *cluster* híbrido entre de melanina-nucleosídeo, que encontrasse localizado entre duas transposases no *contig* 24 (APÊNDICE B).

4.2 Estudo funcional do genoma de *Streptomyces olindensis* DAUFPE 5622

A classificação funcional das proteínas preditas em diferentes grupos ortólogos, utilizando a base de dados de ortólogos não redundantes eggNOG, revelou uma forte ênfase ao metabolismo primário com um 34.5% de regiões codificadoras envolvidas na produção de energia e o metabolismo e transporte de aminoácidos, carboidratos, lipídeos, ácidos nucleicos, coenzimas e íons (Figura 12). Além disto, o genoma de *S. olindensis* utiliza 15.4% para tradução, síntese de ribossomos, transcrição, replicação e reparo do DNA indicando a importância destas funções dentro do genoma de *Streptomyces*. Da mesma forma foi reportado em outros genomas de actinobactérias, onde a maioria de proteínas codificadas no genoma cumpriam funções de manutenção e replicação celular (BENTLEY et al., 2002; SCHWIENTEK et al., 2012; WANG et al., 2013; XU et al., 2014; ZABURANNYI et al., 2014).

Figura 12 Classificação funcional dos CDS em *Streptomyces olindensis* DAUFPE 5622. Distribuição das sequências codificadoras nas categorias do eggNOG. As porções que se encontram fora representam os dois grupos de ortólogos de interesse (A) biossíntese, transporte e catabolismo de metabólitos secundários e (B) proteínas com função desconhecida.

Em relação ao metabolismo secundário, 2.8% das regiões codificadoras foram atribuídas a esta categoria relacionada com o transporte, biossíntese e catabolismo deste tipo de compostos, agrupando 218 proteínas (Tabela 6). Enzimas relacionadas à biossíntese de peptídeos não ribossomais, policetídeos, sideróforos e ectoínas encontram-se neste cluster de ortólogos. Isto concorda com o observado em genomas como *S. coelicolor*, *S. avermitilis*, *S. albus* e *S. bingchengensis* (BENTLEY et al., 2002; IKEDA et al., 2003; WANG et al., 2013; ZABURANNYI et al., 2014), onde o metabolismo secundário representa uma porção significativa do genoma e os genes envolvidos com a produção de enzimas para síntese de NRPS e PKS são ubíquos.

Tabela 6 Funções das proteínas na categoria de biossíntese, transporte e catabolismo de metabólitos secundários

Função Proposta	No. Proteínas	Locus_Tag
Transportador ABC	4	DF19_6769, DF19_6770, DF19_6772, DF19_6771, DF19_7523, DF19_0875, DF19_0929, DF19_5078, DF19_7957, DF19_5078, DF19_7866, DF19_8362, DF19_4816, DF19_8508, DF19_4816, DF19_0929, DF19_3842, DF19_7832, DF19_7833, DF19_7957, DF19_7833, DF19_7832
Dioxigenase	18	DF19_0941, DF19_1127, DF19_1149, DF19_1150, DF19_1378, DF19_2159, DF19_2572, DF19_2868, DF19_3000, DF19_3389, DF19_3397, DF19_3889, DF19_4592, DF19_4824, DF19_4836, DF19_4860, DF19_4867, DF19_4914, DF19_5270, DF19_5289, DF19_5369, DF19_5861, DF19_6081, DF19_6115, DF19_7304, DF19_7338,
Citocromo P-450	26	DF19_3253, DF19_3262, DF19_3263, DF19_3339, DF19_3607, DF19_3608, DF19_3609, DF19_3610, DF19_3635, DF19_3636, DF19_3637, DF19_3638, DF19_4037, DF19_4038, DF19_4965, DF19_4984, DF19_4988, DF19_7477, DF19_7478
Sintase de peptídeos não ribossomais (NRPS)	19	DF19_1984, DF19_3296, DF19_4688, DF19_4942, DF19_5251, DF19_5257, DF19_5261, DF19_5263, DF19_5266, DF19_5936, DF19_6747, DF19_6877, DF19_7781
Policetídeo sintase (PKS)	13	DF19_6121, DF19_7298
Beta ceto-acil sintase	2	DF19_0021, DF19_1407
Amidohidrolase	2	DF19_0029, DF19_3012, DF19_3472, DF19_3473, DF19_4604, DF19_6433, DF19_7614, DF19_8681
DSBA oxidoreductase	8	DF19_0027
Ectoína hidroxilase	1	DF19_0233, DF19_0629, DF19_1015, DF19_1223, DF19_1906, DF19_2049, DF19_3514, DF19_3784, DF19_4707, DF19_5728, DF19_7274
Metiltransferase	11	DF19_0241, DF19_0784, DF19_3632, DF19_4989, DF19_4992, DF19_5248, DF19_5278,
Tioesterase	7	DF19_0401, DF19_1718, DF19_1968, DF19_2088, DF19_2213, DF19_2960, DF19_3086, DF19_3257, DF19_3380, DF19_3628, DF19_3748, DF19_4225, DF19_4853, DF19_4886, DF19_5108, DF19_5819, DF19_6633, DF19_6895, DF19_7399, DF19_7483, DF19_7911, DF19_8186, DF19_8418, DF19_8674, DF19_8680, DF19_8705, DF19_8706
Desidrogenase redutase	27	DF19_0783, DF19_0785, DF19_0786, DF19_0787, DF19_0788
Fenilacetato – CoA ligase e oxigenase	5	DF19_0321, DF19_7935, DF19_0939, DF19_1146, DF19_4959
Dieneleactona hidroxilase	5	DF19_0389, DF19_1938, DF19_2388, DF19_2499, DF19_0976, DF19_3290, DF19_3600, DF19_4687, DF19_4945, DF19_6448, DF19_6748, DF19_7861, DF19_7862, DF19_7909, DF19_8139
AMP-dependente sintase e ligase	15	DF19_0254, DF19_8205
Monoxigenase (NADPH)	2	

Fumarilacetoacetato hidrolase	18	DF19_0879, DF19_2940, DF19_2945, DF19_2947, DF19_2986, DF19_2991, DF19_3621, DF19_3626, DF19_4280, DF19_4558, DF19_4578, DF19_4583, DF19_4907, DF19_5401, DF19_6150, DF19_7528, DF19_7865, DF19_8083
Proteína com domínio Rieske (2Fe-2S)	2	DF19_0110, DF19_5137
5-oxoprolinase	3	DF19_2181, DF19_2182, DF19_7769
Descarboxilase	8	DF19_2309, DF19_2118, DF19_3362, DF19_4157, DF19_4910, DF19_7835, DF19_8150, DF19_8151
Esterase	3	DF19_0828, DF19_1468, DF19_3455
Proteína biossíntese de sideróforos	8	DF19_2226, DF19_2227, DF19_2228, DF19_3129, DF19_8203, DF19_5446, DF19_6199, DF19_7558
Acetiltransferase	4	DF19_4174, DF19_7102, DF19_4174, DF19_7102
Deacilase	2	DF19_0551, DF19_1169
Isomerase	2	DF19_0162, DF19_4106
Oxidases	3	DF19_1587, DF19_5142, DF19_7821

Ademais, proteínas relacionadas com a modificação de produtos naturais como metil, acil, glicosil e acetiltransferases, monooxigenases, desidrogenases, isomerases e oxidases estão presentes nesta classificação, assim como proteínas de transporte deste tipo de moléculas. Estes tipos de genes são necessários para a diversidade estrutural e funcional das moléculas produzidas por actinobactérias, assim como mecanismo de resistência contra a possível toxicidade destas nas células produtoras (GETSIN et al., 2013; SALAS; MÉNDEZ, 2007; WALSH; FISCHBACH, 2010).

De forma significativa, 10.8% das proteínas foram agrupadas na categoria de predição da função geral, mas observando detalhadamente encontram-se enzimas como metil e aciltransferases, hidrolases, desidrogenases, fosfoesterases, dioxigenases, oxidoreductases, epimerases, tioesterases, ciclases e proteínas transportadoras que podem atuar na produção e modificação de produtos naturais, ampliando o número de regiões codificadoras que podem estar associadas a metabolismo secundário.

4.3 Genômica comparativa entre espécies de *Streptomyces*

Utilizando a bases de dados do KEGG para relacionar os genes presentes em *S. olindensis* a diversas vias metabólicas, as proteínas foram agrupadas em grupos de ortólogos, comparando estes clusters com genomas de referência. Neste caso,

foram atribuídos 1223 grupos de ortólogos que são compartilhados pelas cinco espécies, conformando o *core* genoma entre estas bactérias (Figura 13).

Dentro destes grupos, podem-se encontrar proteínas envolvidas em vias metabólicas primárias como o ciclo dos ácidos tricarboxílicos, glicólise e gliconeogênese, via do shikimato, síntese de aminoácidos, ácidos nucleicos e de ácidos graxos. Também é possível identificar proteínas como ATPases envolvidas na geração de energia, proteínas reguladoras, proteínas transportadoras com domínios transmembrana, assim como, proteínas encarregadas da síntese das unidades ribossomais e enzimas de replicação do DNA como polimerases, topoisomerases, helicases e girases.

Estudos de genômica comparativa no *phylum Actinobacteria* e mais especificamente no gênero *Streptomyces* tem demonstrado que o *core* genoma, ou conjunto de genes indispensáveis, está formado principalmente, por proteínas envolvidas em vias de metabolismo primário, assim como manutenção e replicação da célula (ALAM et al., 2011; DOROGHAZI; BUCKLEY, 2014; KIM et al., 2015; ZHOU et al., 2012)

Algumas proteínas incluídas dentro deste *core* genoma participam na biossíntese de metabólitos secundários, alguns parecem ser essenciais para as bactérias deste gênero, como a síntese de ectoína, biossíntese de reguladores pleiotrópicos como as butirolactonas, produção de sideróforos para captação de ferro, síntese de melanina e pigmento de esporos e síntese de hopanoides (DOROGHAZI; BUCKLEY, 2014; ZHOU et al., 2012). Outros genes pelo contrário, codificam para proteínas envolvidas na biossíntese de metabólitos secundários que não são essenciais, mas podem apresentar uma vantagem competitiva como os policetídeos e peptídeos não ribossomais, já que estes estão presentes em todos os genomas de *Streptomyces* sequenciados até o momento (BENTLEY et al., 2002; IKEDA et al., 2003; WANG et al., 2013; ZABURANNYI et al., 2014; ZHOU et al., 2012).

Figura 13 Diagrama de Venn com os grupos de ortólogos da base de dados KEGG de acordo com a informação de vias metabólicas de *S. olindensis*, comparados com os grupos de ortólogos de quatro espécies de *Streptomyces*. Informação obtida do IMG-JGI

A respeito dos elementos únicos no genoma de *S. olindensis*, foram identificados 108 grupos de ortólogos dentro das categorias funcionais das vias metabólicas representadas no KEGG *ortholog database* (Figura 14). Algumas destas categorias estão envolvidas no metabolismo de carboidratos e aminoácidos, assim como síntese de parede celular e processos de tradução e transcrição. É interessante ressaltar que alguns destas categorias pertencem a vias biossintéticas para a produção de moléculas tipo PKS, NRPS, sideróforos e terpenoides, sugerindo que a diversidade do metabolismo secundário é parcialmente responsável pela exclusividade de alguns ortólogos no cromossomo de *S. olindensis* (KIM et al., 2015; ZHOU et al., 2012). Outros grupos únicos encontrados estão compostos por proteínas de transporte como permeases, transportadores tipo ABC e facilitadores MFS, altamente conhecidos por seu desempenho como mecanismo de resistência e

secreção nos clusters biossintéticos (GETSIN et al., 2013). Adicionalmente, como partes destes grupos singulares nesta linhagem estão os reguladores da transcrição dos genes comprometidos na síntese de compostos de interesse biotecnológico (LIU et al., 2013).

Figura 14 Distribuição funcional dos 108 grupos de ortólogos únicos em *Streptomyces olindensis*

4.4 Análise Filogenética de *Streptomyces olindensis* DAUFPE 5622

4.4.1 Filogenia inferida utilizando o gene 16S rDNA

Foi realizada uma árvore filogenética do gene rDNA 16S de *S. olindensis* (Figura 15), determinando que as espécies mais próximas à bactéria em estudo são *S. carpinensis*, *S. purpurascens* e *S. levis*. A informação genômica disponível das duas espécies relacionadas não é suficiente para determinar se há semelhanças além do gene de rDNA 16S, mas observando a distribuição dos genomas das diferentes espécies utilizadas na construção da árvore, as actinobactérias amplamente conhecidas pela produção de diversos metabólitos secundários, encontram-se dispersas pela árvore uniformemente. Isto, possivelmente, devido a que a região codificadora do gene 16S de rDNA é uma sequência altamente conservada devido à baixa pressão seletiva por sua função estrutural na síntese dos

ribossomos em bactérias (JANDA; ABBOTT, 2007); a diferença dos genes responsáveis pelo metabolismo secundário que conferem vantagens evolutivas, a maioria destes são encontrados nos extremos do cromossomo linear no caso de *Streptomyces* sp., sítios sensíveis a deleções e inserções (DOROGHAZI; METCALF, 2013).

Figura 15 Árvore filogenética rDNA 16S de *Streptomyces olindensis* DAUFPE 5622. Árvore construída utilizando o gene 16S rDNA pelo método de máxima verossimilhança e as linhagens *Thermobifida fusca* e *Micromonospora peucetia* DSM 43363 como grupos externos. Valores de bootstrap menores a 60% não representados na figura.

4.4.2 Filogenia concatenada e comparação de diversidade de metabólitos secundários

A partir do observado na primeira árvore filogenética, decidiu-se comparar a *Streptomyces olindensis* com bactérias que possuíam genomas completos depositados nas bases de dados. Utilizaram-se inicialmente os genomas de *Streptomyces* de referencia como *S. coelicolor*, *S. albus*, *S. avermitilis* e *S. bingchenggensis* e adicionalmente actinobactérias produtoras de antraciclinas semelhantes ao policetídeo tipo II Cosmomicina D. Foi realizada uma árvore filogenética utilizando o gene 16S rDNA e quatro genes de *housekeeping* já descritos para a construção de filogenias do gênero (ALAM et al., 2010; VENTURA et al., 2007). Como grupo externo foi utilizado a linhagem da actinobactéria *Micromonospora aurantiaca* ATCC 27029, gênero de bactérias que também produzem moléculas das famílias de antraciclinas (SOUSA et al., 2012). A espécie mais próxima à *S. olindensis* DAUFPE 5622 segundo a filogenia construída é a linhagem *Streptomyces* sp. NRRL WC-3641 seguido de *S. zinciresistens* K42 e *S. coelicolor* A3(2) (Figura 16). Partindo deste contexto filogenético, foi realizada a comparação dos *clusters* biossintéticos putativos nos genomas das espécies na árvore, ao nível de família de produto natural.

Figura 16 Árvore filogenética concatenada de *Streptomyces olindensis* DAUFPE 5622 e diversidade de clusters de metabólitos secundários. Esta árvore foi construída utilizando o gene 16S rDNA e genes *housekeeping* (*rpoB*, *gyrB*, *atpD*, *recA*) pelo método de máxima verossimilhança e a linhagem *Micromonospora aurantiaca* ATCC 27029 como grupo externo. O gráfico de barras representa o número de clusters de metabólitos secundários em cada genoma na mesma ordem da árvore filogenética

Encontrou-se uma correlação positiva entre o tamanho do genoma e o número de vias biossintéticas (DOROGHAZI; METCALF, 2013). Em quanto à diversidade de clusters, é possível afirmar a ubiquidade das famílias de NRPS, PKS, terpenos e híbridos entre eles em todas as espécies na árvore. Este dado é confirmado pela presença de proteínas para produção de hopanóides, carotenoides, pigmento de esporos (PKS), e enzimas multimodulares para produção de peptídeos não ribossomais, que estão envolvidas na biossíntese deste tipo de compostos, dentro do *core* genoma destas bactérias (KIM et al., 2015; ZHOU et al., 2012).

Por outro lado, os lantipeptídeos se encontram na maioria dos genomas estando ausentes só em *S. avermitilis* MA-4680, *Streptomyces* sp. CNT302 e *S. albus* J1074. Outro tipo de metabólitos que apresentam uma vantagem adaptativa como assimilação de substâncias orgânicas e osmoproteção, sideróforos, melanina e reguladores globais como as butirolactonas estão presentes na maioria dos genomas de actinobactérias em especial no gênero *Streptomyces* (SADEGHI et al., 2014).

Como particularidade, exclusivamente em *S. olindensis* foram encontrados dois clusters biossintéticos que pertencem a duas famílias de moléculas bioativas. O primeiro, um antibiótico da família dos aminociclitóis, compostos que possuem atividade antibacteriana, antifúngica e antidiabética (MAHMUD; LEE; FLOSS, 2001; MAHMUD, 2009). O segundo *cluster* biossintético encontrado em *S. olindensis* pertence à família dos fosfonatos, os quais são usados como herbicidas, antimaláricos, antimicrobianos, antifúngicos e anti-hipertensivos (CIONI et al., 2014).

4.5 Identificação do potencial biotecnológico utilizando a abordagem “*genome mining*” e descrição dos *clusters* biossintéticos achados no genoma de *S. olindensis* DAUFPE 5622

A detecção de vias biossintéticas de produtos naturais feita no software antiSMASH revelou a presença de 37 possíveis vias, representando o potencial biotecnológico da linhagem de *Streptomyces olindensis*. Ao realizar uma anotação manual das sequências o número de *clusters* reduziu-se a 33 possíveis *clusters* biossintéticos (Tabela 7), incluindo um *cluster* gênico que pertence à síntese de um *cluster* que codifica para a produção de outro terpeno achado manualmente.

Está análise com o software antiSMASH permitiu comparar o encontrado na análise global usando a agrupação de ortólogos com a base de dados eggNOG,

onde só 2.8% das proteínas do genoma estavam envolvidas na síntese e transporte de metabólitos secundários (Figura 12). Após a identificação e curadoria manual dos *clusters* biossintéticos calculou-se que 13.6% das proteínas do genoma estão dedicadas à biossíntese de produtos naturais, ocupando 1.27 Mb do genoma total. Este fenômeno tem sido observado em outros genomas de actinobactérias, em especial do gênero *Streptomyces* onde a maioria de espécies emprega entre 10% e 20% do genoma para a produção de metabólitos, a maioria de interesse biotecnológico (BENTLEY et al., 2002; WANG et al., 2013; ZABURANNYI et al., 2014).

O servidor web antiSMASH identifica automaticamente potenciais *clusters* biossintéticos integrando o algoritmo probabilístico ClusterFinder (CIMERMANCIC et al., 2014), que auxilia na detecção de *clusters* gênicos de famílias desconhecidas, utilizando modelos ocultos de Markov para criar perfis HMM de enzimas comumente encontradas em *clusters* biossintéticos como oxidoreductases e metiltransferases (WEBER et al., 2015), e o módulo ClusterBlast que identifica *clusters* semelhantes a aqueles previamente depositados no *MIBiG community project* (MEDEMA, 2015) como ferramenta de de-replicação.

Adicionalmente o software prediz os sítios ativos das enzimas chaves, assim, como as unidades incorporadas no caso de policetídeos e peptídeos não ribossomais. Estas análises preliminares auxiliam na anotação manual dos genes envolvidos e na elucidação da possível estrutura química do produto final (WEBER et al., 2015).

Tabela 7 *Clusters* biossintéticos detectados pelo software antiSMASH dentro do genoma de *Streptomyces olindensis*

Localização <i>cluster</i> gênico	Tipo de Produto Natural	Produto predito
Conhecidos		
DF19_RS02990 - DF19_RS03200	Policetídeo Tipo II	Cosmomicina D
Preditos		
DF19_RS00100 - DF19_RS00155	Exopolissacarídeo	5-Hidroxiectoina
DF19_RS00335 - DF19_RS00405	Terpeno	Desconhecido
DF19_RS02435 - DF19_RS02580	Aminociclitol	Desconhecido
DF19_RS05010 - DF19_RS05030	Terpeno	2-Metilisoborneol
DF19_RS05010 - DF19_RS05075	Melanina	Desconhecido
DF19_RS07610 - DF19_RS07705	Terpeno	Geosmina - Gemacradienol

DF19_RS08415 - DF19_RS08505	RiPP - Linaridina	Desconhecido
DF19_RS10405 - DF19_RS10440	Butirolactona	Desconhecido
DF19_RS10710 - DF19_RS10760	Sideróforo	Desconhecido
DF19_RS13360 - DF19_RS13450	Fosfonato	Fosacetamicina
DF19_RS15025 - DF19_RS15065	Sideróforo	Desconhecido
DF19_RS15565 - DF19_RS15765	NRPS	Desconhecido
DF19_RS15770 - DF19_RS16145	Outros	Desconhecido
DF19_RS17195 - DF19_RS17360	Sideróforo NRPS	Amiquelina-like
DF19_RS17370 - DF19_RS17545	NRPS-PKSI	Desconhecido
DF19_RS41950 - DF19_RS17910	Butirolactona	Desconhecido
DF19_RS18615 - DF19_RS18655	Butirolactona	Desconhecido
DF19_RS19275 - DF19_RS19495	NRPS-PKSI	Desconhecido
DF19_RS22390 - DF19_RS22585	PKSI	Desconhecido
DF19_RS22880 - DF19_RS22990	RiPP - Lantipeptídeo	Desconhecido
DF19_RS23585 - DF19_RS23965	NRPS-PKSI	Desconhecido
DF19_RS25065 - DF19_RS25240 DF19_RS32375 - DF19_RS32390	PKSI	Tetronomicina-like
DF19_RS27645 - DF19_RS27790	Melanina- Nucleosídeo	Desconhecido
DF19_RS28095 - DF19_RS28150	Terpeno	Albaflavenona
DF19_RS29250 - DF19_RS29450	PKSII	Pigmento de esporos
DF19_RS33020 - DF19_RS33050	Terpeno	Desconhecido
DF19_RS35720 - DF19_RS35930	NRPS-PKSI	Desconhecido
DF19_RS36550 - DF19_RS36620	RiPP - Lasso-peptídeo	Desconhecido
DF19_RS37190 - DF19_RS37255	RiPP - Linaridina	Desconhecido
DF19_RS39200 - DF19_RS39295	Sideróforo	Desferroxamina
DF19_RS41160 - DF19_RS41280	Terpeno	Hopanoides
DF19_RS16240 - DF19_RS16280	Terpeno	Desconhecido

4.5.1 Descrição dos genes associados à produção de γ -Butirolactona

As gama butirolactonas são moléculas produzidas pela maioria de *Streptomyces* e outras actinobactérias, e estão envolvidas na ativação do metabolismo secundário. A butirolactona mais estudada é o Fator-A (2-isocaprilóil-3R-hidroximetil- γ -butirolactona) de *Streptomyces griseus*. Este Fator-A é sintetizado pelo gene *afsA* para regular a produção de estreptomicina, grixazona e iniciar a diferenciação morfológica na bactéria (BIBB, 2005).

No genoma de *S. olindensis* evidenciou-se a presença de três *clusters* biossintéticos para a produção de butirolactonas, que poderiam atuar como reguladores pleiotrópicos do metabolismo primário e secundário de *Streptomyces*

encarregadas de ativar a esporulação, a produção de micélio aéreo e a síntese alguns metabólitos secundários (Figura 17).

Figura 17 Genes biossintéticos para a produção de γ -butirolactonas presentes em *S. olindensis* DAUFPE 5622

Dentro dos três *clusters* biossintéticos está o homólogo ao gene encarregado da síntese do Fator-A ou gama butirolactona denominado *afsA* (Figura 17). Este gene codifica para uma enzima que condensa uma cadeia ceto acil e um fosfato dihidroxiacetona para formar o Fator-A (KATO et al., 2007). A união desta molécula à proteína ArpA no citoplasma, sintetizada pelo gene *arpA* libera a transcrição da proteína AdpA. Está cascata de regulação é necessária para a síntese do antibiótico aminoglicosídeo estreptomicina já que a presença da proteína AdpA ativa o regulador específico (*strR*) do *cluster* em *S. griseus* (BIBB, 2005).

Embora nas três vias biossintéticas achou-se um gene homologa ao *afsA*, só numa (*Cluster 3*) foi achado o homólogo ao gene *arpA* que se encontra contíguo no *cluster* (Figura 17). A presença deste gene do lado do *afsA* é característico de *clusters* que codificam para butirolactonas que podem estar envolvidas só na ativação do metabolismo secundário e não em diferenciação morfológica (BIBB, 2005; MARTÍN; LIRAS, 2010). Além disso, existem ORFs que poderiam estar envolvidas nas modificações da molécula como desidrogenases, epoxidases, fosfatases e monooxigenases contíguas a ORF *asfA-like* (KATO et al., 2007).

Adicionalmente, nos *clusters* de biossíntese de moléculas reguladoras observam-se alguns genes que codificam para reguladores transcricionais da família TetR, ou componentes como sensores de histidina que atuam como reguladores de vias biossintéticas. A proximidade destes homólogos aos genes de síntese de butirolactonas pode indicar a ação destas moléculas como ativadores de reguladores específicos de vias para a biossíntese de produtos naturais (ZOU et al., 2014).

4.5.2 Anotação da via biossintética para produção de ectoína

A ectoína e seus derivados pertencem ao grupo dos metabólitos secundários que são produzidos por um grande número de bactérias incluindo o gênero *Streptomyces*. As proteínas responsáveis pela síntese deste metabólito são encontradas dentro do *core* genoma destas bactérias, indicando a importância dele. Este metabólito é exportado fazendo parte do grupo de exopolissacarídeos que tem não só como função a osmoproteção da célula, mas também afetam a estabilidade e a dobragem das proteínas e protegem o DNA das condições de estresse (BURSY et al., 2008; SADEGHI et al., 2014).

A sequência do DNA de *S. olindensis* revelou a presença de três ORFs que codificam para três enzimas, uma diaminobutirato acetiltransferase, a diaminobutirato piruvato aminotransferase e a ectoína sintase, proteínas necessárias para a síntese de ectoína (operon *ectABC*) (BURSY et al., 2008; GAN et al., 2013; SADEGHI et al., 2014) e adicionalmente, a presença da ORF que codifica para a ectoína hidroxilase *ectD* para a produção do derivado hidroxietoína (Tabela 8) (BURSY et al., 2008), todas, orientadas na mesma posição (Figura 18).

Figura 18 *Cluster* gênico para produção de hidroxietoína presente em *Streptomyces olindensis*

Também, foram encontradas quatro ORFs (*ectI*, *ectII*, *ectIII* e *ectIV*) envolvidas na conversão de asparagina a L-aspartato- β -semialdeído como precursor da síntese de ectoína (BURSY et al., 2008b). Dentro da agrupação gênica acharam-se quatro ORFs adicionais que codificam para uma aminotransferase, uma oxidoreductase, uma proteína possível reguladora e uma proteína hipotética (Tabela 8), mas ainda possuem uma função desconhecida na produção de hidroxiectoína.

Tabela 8 Funções preditas das ORFs no cluster biossintético de hidroxiectoína

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
<i>ectII</i>	363	Asparagina Amidohidrolase- Aspartato amidohidrolase	95/98	<i>S. viridochromogenes</i> WP_003989359
<i>ectIII</i>	366	Aspartato aminotransferase	92/96	<i>S. viridochromogenes</i> WP_039932170
<i>ectVIII</i>	177	Proteína hipotética	84/87	<i>S. viridochromogenes</i> WP_003989361
<i>ectA</i>	154	Diaminobutirato Acetiltransferase	97/98	<i>S. viridochromogenes</i> WP_003989362
<i>ectB</i>	423	Diaminobutirato-piruvato aminotransferase	96/98	<i>S. afghaniensis</i> WP_020277269
<i>ectC</i>	132	L-ectoína sintase	95/98	<i>S. afghaniensis</i> WP_020277268
<i>ectD</i>	302	Ectoína hidroxilase	96/97	<i>S. afghaniensis</i> WP_020277267
<i>ectV</i>	350	Aminotransferase	90/95	<i>S. viridochromogenes</i> WP_039932171
<i>ectVI</i>	240	DSBA oxidoreductase	93/95	<i>S. afghaniensis</i> WP_010037844
<i>ectI</i>	159	Acetiltransferase (N-acetyl-L-aspartato)	80/84	<i>S. afghaniensis</i> WP_037673963
<i>ectVII</i>	202	Regulação hipotética da enolase	95/96	<i>S. afghaniensis</i> WP_020277263
<i>ectIV</i>	509	Semialdeído desidrogenase	98/99	<i>S. chartreusis</i> WP_010037850

4.5.3 Vias biossintéticas para produção de pigmentos

A produção de pigmentos em *Streptomyces* tem sido amplamente estudada devido às diferenças morfológicas macroscópicas deste gênero, e está dividida em vários tipos de *clusters* biossintéticos (FUNA et al., 2005). A produção de melanina derivada de tirosina e o pigmento de esporos da família policetídica, encontram-se claramente apresentados no genoma da bactéria em estudo e é possível observar a sua produção durante o crescimento de *S. olindensis*.

4.5.3.1 Cluster Melanina

A melanina é um pigmento marrom escuro de alto peso molecular produzido abundantemente no gênero *Streptomyces* e a produção deste composto pelas bactérias, tem sido usada como uma ferramenta de classificação taxonômica. Estudos tem demonstrado que a presença deste composto nas células aumenta a resistência a radiação UV, tendo um papel fundamental na proteção contra estresse ambiental (FUNA et al., 2005). Além disso, atua na absorção de radicais livres e sequestro de ferro (GUO et al., 2014). Estas moléculas têm se tornado interessante pela atividade biológica como antioxidante, antimicrobiana e antitumoral (SIVAPERUMAL et al., 2014).

O intermediário principal para a produção de melanina em actinobactérias é a L-dihidroxifenilalanina (L-DOPA), que é sintetizada pela oxidação do aminoácido tirosina usando como catalizador a enzima tirosinase (MENCHER; HEIM, 1962). A partir deste precursor, oxidações não enzimáticas ocorrem para sintetizar melanina (GUO et al., 2014). Na Figura 19 observa-se a presença da enzima tirosinase e o cofator codificados pelas ORF *melC1* e *melC2* e adicionalmente 5 ORFs possivelmente envolvidas na modificação posterior de moléculas como monooxigenases, desidrogenases e redutases (Tabela 9).

Figura 19 *Cluster* gênico para produção de melanina presente em *Streptomyces olindensis*

Adicionalmente, dentro do cluster biossintético encontraram-se três ORFs possivelmente envolvidas no transporte da molécula (*orfG*, *orfH*, *orfJ*) que apresentam homologia com transportadores de aminoácidos ou moléculas derivadas deles (Tabela 9).

Tabela 9 Funções preditas das ORFs no cluster biossintético da melanina

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
<i>orfA</i>	233	Ribonuclease HI	95/96	<i>S. afghaniensis</i> WP_020269725
<i>melC2</i>	273	Tirosinase	95/97	<i>S. iakyrus</i> WP_033309423
<i>melC1</i>	122	Cofator Tirosinase <i>melC</i>	88/90	<i>S. viridochromogenes</i> WP_039931758
<i>orfB</i>	89	Monooxigenase dependente de FAD	95/100	<i>S. leeuwenhoekii</i> WP_029387575
<i>orfD</i>	579	Desidrogenase dependente de FAD/FMN	89/92	<i>S. chartreusis</i> WP_029180713
<i>orfE</i>	327	Aldo-ceto reductase	93/96	<i>S. cellulosa</i> WP_030676174
<i>orfF</i>	116	Proteína hipotética (Prolina reductase)	91/91	<i>S. chartreusis</i> WP_010033159
<i>orfG</i>	141	Transportador Arginina/ornitina ArcD	96/98	<i>S. chartreusis</i> WP_029180714
<i>orfH</i>	141	Proteína de membrana	91/95	<i>S. viridochromogenes</i> WP_003987910
<i>orfJ</i>	346	Proteína de membrana	96/97	<i>S. iakyrus</i> WP_051814706

4.5.3.2 Pigmento de esporos (Policetídeo Tipo II)

Na maioria das bactérias pertencentes ao gênero *Streptomyces*, identifica-se uma via biossintética do tipo PKS II para a produção de um pigmento característico dos esporos, a exceção da bactéria *Streptomyces griseus* que possui um PKS tipo III para a produção do pigmento característico de esporos (FUNA et al., 2005). A produção desta molécula é ativada quando o processo de diferenciação morfológica começa a formação de micélio aéreo (NOVAKOVA; BISTAKOVA; KORMANEC, 2004; YU; HOPWOOD, 1995). Em *Streptomyces coelicolor* a produção deste pigmento foi caracterizada levando a identificação de oito ORFs mínimas para a síntese da molécula (*whiE*-ORFI ate *whiE*-ORFVIII).

Em *Streptomyces olindensis* foram achadas nove ORFs homólogas as encontradas em *S. coelicolor* para a síntese básica do pigmento (Figura 20), desde a síntese do policetídeo mínimo (*whiEIII*, *whiEIV* e *whiEV*), ORFs envolvidas na ciclização da molécula como ciclases e aromatasas (*whiEII*, *whiEVI* e *whiEVII*), ORFs encarregadas do direcionamento da molécula dentro dos esporos (*whiEI*) e por ultimo uma ORF que atua como hidroxilase do produto final, mudando sua coloração (YU; HOPWOOD, 1995). Adicionalmente duas ORFs representando potenciais reguladores transcricionais que podem estar envolvidas na ativação da

via (*orf1*, *orf4*) (Tabela 10) e quatro ORFs hipotéticas que podem realizar passos de hidrólise e oxidação na via biossintética.

Figura 20 *Cluster* gênico para produção de pigmento associado aos esporos presente em *Streptomyces olindensis*

Tabela 10 Funções previstas das ORFs no *cluster* biossintético para produção de pigmento de esporos. Percentagem de identidade com respeito aos homólogos achados em *S. coelicolor*

ORF	No. a.a.	Função Proposta	%Ident./ Simil.	Homólogo e Origem No. Acesso	% Ident. <i>S. coelicolor</i>
<i>orf1</i>	219	Regulador transcricional TetR	93/94	<i>S. viridochromogenes</i> EFL34939	88% NP_629557
<i>orf2</i>	566	Isobutiril coA mutase alfa	98/98	<i>S. afghaniensis</i> WP_020276087	94% NP_629554
<i>orf3</i>	108	Proteína de membrana	94/97	<i>S. afghaniensis</i> WP_020276088	81% NP_629553
<i>orf4</i>	169	Regulador transcricional MarR	98/99	<i>S. chartreusis</i> WP_010048577	94% NP_629552
<i>whiE I</i>	88	Hipotética monooxygenase	97/98	<i>S. afghaniensis</i> WP_037672446	41% NP_631499
<i>orf5</i>	322	Citocromo P450-hidroxilase	97/98	<i>S. chartreusis</i> WP_010048575	45% NP_627960
<i>whiE VIII</i>	508	Policetídeo hidroxilase	94/96	<i>S. afghaniensis</i> WP_037672447	67% NP_629463
<i>orf6</i>	327	Hidrolase	95/97	<i>S. afghaniensis</i> WP_037672448	
<i>whiE VII</i>	113	Policetídeo ciclase	76/86	<i>S. cyaneogriseus</i> WP_044378734	71% NP_629456
<i>whiE VI</i>	159	Aromatase	96/96	<i>S. afghaniensis</i> WP_020276096	74% NP_629457
<i>whiE V</i>	85	Proteína carreadora de acil ACP	93/94	<i>S. afghaniensis</i> WP_020276097	61% NP_629458
<i>whiE IV</i>	401	Ceto acil sintase subunidade beta (CLF)	95/96	<i>S. afghaniensis</i> WP_020276098	74% NP_629459
<i>whiE III</i>	422	Ceto acil sintase subunidade alfa	97/98	<i>S. chartreusis</i> WP_010048567	78% NP_629460
<i>whiE II</i>	154	Policetídeo ciclase	98/98	<i>S. chartreusis</i> WP_010048566	70% NP_629461
<i>whiE I</i>	372	Hipotética monooxygenase	96/98	<i>S. afghaniensis</i> WP_020276101	62% NP_629462

De acordo com as características microscópicas de *S. olindensis* a bactéria expressa este cluster gênico no momento de diferenciação morfológica e produção de micélio aéreo para assim levar aos esporos.

4.5.4 Cluster gênico da biossíntese do fosfonato em *S. olindensis*

Os fosfonatos são moléculas amplamente empregadas como inibidores de enzimas que reconhecem substratos fosforilados. Os enlaces característicos C-P e C-P-C e sua similaridade geométrica com o éster fosfato fazem destes tipos de compostos candidatos ideais para ser usados com “imitadores” de metabólitos intracelulares (ENGEL, 1977; EVANS et al., 2013; METCALF; VAN DER DONK, 2009). Dentro dos *clusters* biossintéticos destas moléculas a principal característica é a presença da fosfoenolpiruvato fosfomutase (PepM), fato que tem incrementado a descoberta de novas biomoléculas utilizando a abordagem de *genome mining* (JU; DOROGHAZI; METCALF, 2014).

No genoma de *S. olindensis* foi achado um possível *cluster* biossintético para a produção de um fosfonato (Figura 21). Esta agrupação gênica demonstrou homologia com a via biossintética da rizotocina (BORISOVA et al., 2010), fosfonato que possui atividade antifúngica. Após a anotação detalhada do *cluster* achou-se homologia com a via para produção de fosacetamicina (EVANS et al., 2013), composto análogo a rizotocina, mas que possui um grupo acetamida adicional, envolvendo genes biossintéticos a mais (Tabela 11).

Figura 21 Agrupação gênica para a produção da molécula da família dos fosfonatos

Tabela 11 ORFs preditas no *cluster* biossintético de fosfonato

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
phoS	1294	Regulador transcricional SARP	84/89	<i>S. afghaniensis</i> EPJ39138
phoSQ	297	Glicosiltransferase Tipo II	91/94	<i>S. cyaneogriseus</i> WP_044382377
phoS	106	ATPase	41/51	<i>Singulisphaera acidiphila</i> WP_015246538
phoS	320	Permease –transportador de metabólitos/peptídeos	89/91	<i>S. davawensis</i> WP_015658823
phoS1	165	Proteína hipotética	98/98	<i>S. davawensis</i> WP_015658824
phoS2	819	Proteína hipotética (Domínio ceto-bisfosfato aldolase)	88/91	<i>Streptomyces aureus</i> WP_051900845
phoS3	481	Glicina/Serina hidroximetiltransferase	26/38	<i>S. scabiei</i> WP_013000533
phoS4	232	Proteína hipotética	27/37	<i>Actinoplanes</i> sp. SE50/110 WP_014692816
phoS	108	Fosfonopiruvato descarboxilase cadeia alfa <i>ppd</i>	94/95	<i>Streptomyces aureus</i> AGH16366
phoS	190	Fosfonopiruvato descarboxilase	87/92	<i>Streptomyces aureus</i> AGH16367
phoS	341	4-hydroxy-2-oxovalerate aldolase	30/46	<i>S. davawensis</i> WP_015656163
phoS	310	Fosfoenolpiruvato fosfomutase <i>pepM</i>	94/96	<i>S. davawensis</i> WP_015658830
phoS	371	Desidratase	89/93	<i>Streptomyces aureus</i> AGH16370
phoS5	154	Tiamina fosfato sintase	95/96	<i>S. davawensis</i> WP_015658832
phoS	380	Fosfato aminotransferase	98/99	<i>S. davawensis</i> WP_015658833
phoS	101	Glutaredoxina	87/94	<i>Streptomyces aureus</i> AGH16373
phoS	400	ATPase <i>grasp</i> (carbamoyl fosfato sintetase)	37/52	<i>S. vietnamensis</i> WP_041130659
phoS	226	Fosfodiesterase	87/91	<i>S. davawensis</i> WP_015658836
phoS	385	ATPase <i>grasp</i> (carboxilase)	42/54	<i>S. griseoaurantiacus</i> WP_006138524
phoS	437	Proteína de Transporte (MFS)	86/89	<i>S. davawensis</i> WP_015658838

4.5.5 Biossíntese de Terpenos em *S. olindensis* DAUFPE 5622

Os terpenos são um dos produtos naturais mais abundantes na natureza descritos até hoje, curiosamente, a maioria tem sido isolada de plantas e fungos e na atualidade, poucas estruturas identificadas de bactérias (CANE; IKEDA, 2012). Com a disponibilização da informação genômica, tem-se identificado uma grande quantidade de terpeno sintases (enzima principal), especialmente em actinobactérias do gênero *Streptomyces* e cianobactérias do gênero *Nostoc* spp (NAKANO; HORINOUCI; OHNISHI, 2011).

A identificação dos motivos característicos dentro das terpeno sintases tem sido um grande desafio, devido à baixa similaridade com as enzimas identificadas em eucariotas, mas tipicamente demonstram dois domínios altamente conservados de união ao Mg^{2+} : O primeiro um motivo com alta quantidade de aspartato e o segundo uma tríade catalítica NSE/DTE (CANE; IKEDA, 2012). A ampla diversidade estrutural das moléculas terpenoides deve-se a dois fatores, primeiro a especificidade da enzima para escolher o precursor utilizado na síntese, o qual pode ser geranyl difosfato (GPP - C_{10}), farnesil difosfato (FPP - C_{15}) e gerani-geranyl difosfato (GGPP - C_{20}) e segundo, as modificações enzimáticas posteriores como metilações, oxidações ou adições de cadeias policetídicas (CANE; IKEDA, 2012; NAKANO; HORINOCHI; OHNISHI, 2011).

O software antiSMASH, utiliza perfis HMM usando diferentes terpeno sintases típicas que contém estes motivos para assim identificar *in-silico* possíveis ORFs homólogas, aumentando o número de terpeno sintases identificadas nos genomas sequenciados.

4.5.5.1 Biossíntese de terpenos caracterizados

No genoma de *S. olindensis* foram identificados 6 possíveis *clusters* biossintéticos para a síntese de terpenos pelo antiSMASH e adicionalmente, uma via anotada manualmente. Quatro destes *clusters* mostraram homologia com proteínas já descritas para síntese dos terpenos voláteis geosmina, 2-metilisoborneol e albaflavenona (JIANG; HE; CANE, 2006; ZHAO et al., 2008), e o quarto apresenta homologia com o *cluster* para a produção de hopanóides, estruturas terpenóides que tem como função biológica proporcionar estabilidade permeabilidade à membrana lipídica celular nas bactérias (Figura 22) (PAN et al., 2015).

A presença da enzima terpeno sintase ou ciclase é a principal característica destes *clusters* biossintéticos, já que são as encarregadas de ciclar os precursores (CANE; IKEDA, 2012), seguido da presença em alguns casos (síntese de albaflavenona) de uma monooxigenase ou citocromo p450 que está envolvida em rearranjos oxidativos (YAMADA et al., 2015). Dependendo do tipo de composto outros genes específicos estão envolvidos na modificação antes e após a síntese do produto como é o caso do 2-metilisoborneol (Figura 22), monoterpene volátil que é gerado a partir do precursor geranyl difosfato. Este é modificado previamente por

uma enzima metiltransferase que no caso de *S. olindensis* é a *orfB* (APÊNDICE B). Adicionalmente os *clusters* achados apresentam geralmente genes reguladores encarregados da ativação ou repressão das vias metabólicas, e ORFs envolvidas no transporte da molécula, dependendo da sua função biológica dentro da célula.

No caso da via biossintética dos hopanóides, acharam-se ORFs homólogas para a produção de hopeno a partir de farnesil difosfato, passando pelo precursor de isoprenóides, esqualeno (APÊNDICE B). Estudos que demonstram que a produção de hopanóides está relacionada com a diferenciação morfológica, têm associado estas moléculas com a estabilidade e permeabilidade da membrana, especialmente em condições adversas como, mudanças no pH, T°, radiação e salinidade (PAN et al., 2015; PORALLA; MUTH; HÄRTNER, 2000). Isto concorda com as condições que as bactérias do gênero *Streptomyces*, encontradas comumente no solo, enfrentam, levando ao desenvolvimento de mecanismos de defesa ou adaptação.

Figura 22 *Clusters* biossintéticos para a produção de geosmina, 2-metilisoborneol, albaflavenona e hopanóides.

4.5.5.2 Biossíntese de terpenos desconhecidos

Dentro do genoma de *S. olindensis* foram encontrados três potenciais clusters biossintéticos para a síntese de terpenos, sem produtos homólogos conhecidos. Dentro destes foram achadas ORFs homólogas que codificam para terpeno ciclases que são as principais enzimas na síntese de moléculas terpenóides (Figura 23).

Adicionalmente acharam-se monooxigenases encarregadas de oxidar a molécula final, e enzimas como desidrogenases e metiltransferases envolvidas na modificação do produto. Nos clusters detectados observa-se a presença de ORFs reguladoras e transportadoras das moléculas, assim, como proteínas envolvidas na resposta ao estresse (*orfX*, *orfY* do Terpeno 1). Segundo a análise, o cluster do terpeno 1 codifica para um monoterpreno, sendo o precursor deste o geranyl difosfato. O cluster do terpeno 2 possui uma enzima com homologia as Z-isoprenil difosfato sintases, enzimas que possuem um rango catalítico maior sintetizando terpenóides de mais de 15 carbonos (C₁₅) (WANG; OHNUMA, 2000). Por ultimo, o cluster que codifica para o terpeno três, possui uma sesquiterpeno sintase como enzima ciclizadora e duas monooxigenases que podem estar envolvidas na oxidação do produto final (APÊNDICE B).

Figura 23 *Clusters* biossintéticos para a produção de terpenos sem produto conhecido em *S. olindensis*

4.5.6 Peptídeos Ribossomais RiPPs

Os peptídeos de síntese ribossomal (RiPPs) estão divididos em mais de 20 subgrupos, diversidade que confere à família uma ampla diversidade estrutural e biológica, além, das modificações pós-traducionais que estes metabólitos sofrem sendo ubíquos em todas as formas de vida (SCHMIDT, 2010). Em *S. olindensis*

foram encontrados quatro clusters biossintéticos da família dos RiPPs, distribuídos em três subgrupos: os lantipeptídeos, os lassopeptídeos e as linaridinas. Como característica em comum da família, a síntese está mediada por duas regiões que codificam para os peptídeos líder e peptídeo principal, e as diferenças entre os subgrupos dependerão dos aminoácidos incorporados e as modificações pós-traducionais que ocorrem na síntese (ARNISON et al., 2013). Caracterizou-se o cluster do Lantipeptídeo como mecanismo para entender a biossíntese dos peptídeos ribossomais RiPPs. Os clusters biossintéticos do lassopeptídeos e linaridinas estão descritos no APÊNDICE B.

4.5.6.1 Lantipeptídeo Classe III

O software antiSMASH detecta a presença do possível cluster de lantipeptídeos utilizando perfis HMM das enzimas biossintéticas que dispõem os resíduos de aminoácidos nas modificações pós-traducionais.

No caso dos lantipeptídeos, estes possuem resíduos de *meso*-lantionina (Lan) e 3-metillantionina (MeLan), sendo esta a sua característica principal. A lantionina consiste em dois resíduos de alanina com ligações cruzadas via tioéter conectando os β -carbonos (ARNISON et al., 2013). A inserção dos resíduos de Lan e MeLan começa com a desidratação dos resíduos de serina (Ser) e treonina (Thr), tornando-se em dehidroalanina (Dha) e dehidrobutirina (Dhb), respectivamente, seguido da adição de resíduos de cisteína (Cys) (WANG; VAN DER DONK, 2012).

Estes peptídeos são classificados de acordo as enzimas utilizadas para inserir os resíduos Lan e MeLan e no caso de *S. olindensis* foi classificado como lantipeptídeo da classe III, já que a enzima que dispõe os aminoácidos (LanKC) carece de resíduos de união a metais, característicos em lantipeptídeos das classes I, II e IV (ARNISON et al., 2013).

No *cluster* foram identificadas 14 potenciais ORFs envolvidas no metabolismo do lantipeptídeo, sendo característica a presença do homólogo a enzima LanKC e o peptídeo precursor (*lanA*). Adicionalmente, possui quatro ORFs envolvidas no transporte da molécula, duas tendo homologia com proteínas transportadoras de peptídeos ribossomais (*orfV* e *orfW*) (Figura 24).

Figura 24 Cluster gênico e sequência do peptídeo precursor do lantipeptídeo classe III

O peptídeo precursor está dividido na sequência líder que é clivada por uma protease, neste caso a *orfP* (Tabela 12), e a sequência principal que será modificada gerando o peptídeo maduro, quem terá a função biológica.

Tabela 12 Funções previstas das ORFs no cluster biossintético para produção do antibiótico Lantipeptídeo

ORF	No. a.a.	Função Proposta	%Ident/ Simil.	Homólogo e Origem No. Acesso
<i>orf5</i>	139	Inosine monofosfato oxidoreductase	91/95	<i>S. chartreusis</i> WP_010040498
<i>orfS</i>	817	Regulador Serina fosfatase RsbU, subunidade sigma	92/95	<i>S. chartreusis</i> WP_010040502
<i>lanKC</i>	864	Serina treonina quinase + Domínio Lantionina sintase (Ciclase) LanKC	93/95	<i>S. afghaniensis</i> WP_020272067
<i>lanA</i>	38	Proteína classe III lantipeptídeo precursor	97/100	<i>S. viridochromogenes</i> WP_039934377
<i>orfT</i>	583	Proteína transportadora ABC – ATP binding	27/43	<i>Corynebacterium riegellii</i> WP_052205904
<i>orfU</i>	675	Proteína transportadora ABC – ATP binding	81/82	<i>S. chartreusis</i> WP_010040507
<i>orf1</i>	69	Proteína hipotética	77/80	<i>S. afghaniensis</i> EPJ41179
<i>orfR</i>	715	Regulador Serina fosfatase RsbU, subunidade sigma	94/96	<i>S. chartreusis</i> WP_040903315
<i>orf2</i>	141	Proteína hipotética	30/34	<i>Kitasatospora cheerisanensis</i> WP_051653722
<i>orfP</i>	434	Serina Protease	97/99	<i>S. viridochromogenes</i> WP_003994899
<i>orf3</i>	71	Proteína hipotética	99/100	<i>S. chartreusis</i> WP_010040515
<i>orf4</i>	268	Enzima dependente de biotina	93/97	<i>S. afghaniensis</i> WP_020270761
<i>orfV</i>	741	Proteína transportadora de bacteriocina ATPase	97/98	<i>S. viridochromogenes</i> WP_003994896
<i>orfW</i>	941	Proteína transportadora de bacteriocina ATPase	97/98	<i>S. afghaniensis</i> WP_020270759

4.5.7 Anotação da via metabólica do peptídeo não ribossomal NRPS

Os peptídeos de síntese não ribossomal são uma das famílias mais estudadas de produtos naturais, devido a que suas enzimas de síntese são multimodulares e a sua vez possuem vários domínios catalíticos. A variedade estrutural desta família é ampla, já que cada módulo das NRP sintases, pode incorporar diversos aminoácidos que podem sofrer modificações como metilações, acetilações, oxidações e reduções (FELNAGLE et al., 2008). Em *S. olindensis* foram encontrados dois *clusters* de NRPS além de quatro *clusters* híbridos que contêm ORFs associadas à síntese de peptídeos não ribossomais. O cluster homólogo a via da amiquelina foi caracterizado para entender o mecanismo de síntese e incorporação dos aminoácidos na cadeia peptídica.

4.5.7.1 Amiquelina-like

O *cluster* da amiquelina-like contém 30 genes implicados na síntese (Figura 25). Geralmente, estas moléculas são sintetizadas pelas enzimas modulares que integram aminoácidos em cada adenilação e condensação, com a ajuda dos módulos carregadores dos peptídeos, começando com um módulo de adenilação responsável pela formação do aminoacil-AMP (FELNAGLE et al., 2008), que no *cluster* é a *ORF solinE* homóloga à Triostina sintase (PRASEUTH et al., 2008).

Posteriormente, as quatro enzimas multimodulares NRPS codificadas nos genes *solinA*, *solinB*, *solinC* e *solinD* estão encarregadas das condensações de diversos aminoácidos dentro do composto, e estes vão se ligando ao ácido quinoxaline-2-carboxílico que provém do triptofano sintetizado pela antranilato ou aminobenzoato sintase (SCHMOOCK et al., 2005), neste *cluster* a *ORF solinF*. Para a síntese, modificação e adição de aminoácidos como unidades iniciadoras e de alongação da molécula, o *cluster* possui 5 possíveis ORFs responsáveis (*solinH*, *solinJ*, *solinK*, *solinL* e *solinM*). Estes genes estariam envolvidos na síntese e incorporação destes aminoácidos segundo o software antiSMASH: alanina, glutamina, valina, serina, ornitina e cisteína; como unidades extensoras.

Figura 25 Organização do *cluster* gênico envolvido na biossíntese de um peptídeo não ribossomal (NRPS)

Devido à variedade de aminoácidos e modificações que podem aparecer na natureza, os peptídeos não ribossomais, possuem uma grande variedade estrutural e biológica, variando entre cada via biossintética o número e tipo de genes implicados na adição de aminoácidos na molécula (NETT; IKEDA; MOORE, 2009).

A via contém quatro *ORFs* relacionadas com o transporte da molécula fora da célula como mecanismo de resistência que são *solinT1*, *solinT2*, *solinT3* e *solinT4*; além de uma *ORF* envolvida no reparo de DNA como possível mecanismo de resistência (*solinN*). A regulação da via biossintética encontra-se provavelmente mediada por dois reguladores: *solinR1* que é um repressor quinase (ROK) e um ativador transcricional do grupo LysR nomeado *solinR2*. O *cluster* de NRPS contém diversos genes que indicam modificações posteriores como adição de açúcares e a modificação dos mesmos além de um regulador repressor do metabolismo destes açúcares da família DeoR (*solinG1*, *solinG2*, *solinG3* e *solinG4*).

Tabela 13 Funções preditas do *cluster* biossintético para a produção do peptídeo não ribossomal (NRPS) Amiquelina-like

Nome da ORF	No. a.a.	Função Proposta	Ident. / Sim. [%]	Homólogo e Origem
<i>solinO</i>	313	Proteína hipotética secretada (Domínio fosfolipase)	81/90	WP_004003296.1 <i>S. viridochromogenes</i>
<i>solinP</i>	295	Fosfatase (domínio PAP2-like fosfatases e haloperoxidases)	94/94	WP_020276665.1 <i>S. afghaniensis</i>

<i>solinT1</i>	549	Transportador ABC	57/71	WP_020872360.1 <i>S. rapamycinicus</i>
<i>solinT2</i>	355	Transportador ABC	57/73	WP_019744684.1 <i>Rhodococcus erythropolis</i>
<i>solinE</i>	642	Triostina Sintase (Domínio de Adenilação)	99/99	WP_031120685.1 <i>Streptomyces</i> sp. NRRL WC-3641
<i>solinF</i>	497	Antranilato sintase subunidade I - Aminobenzoato sintase	29/42	WP_019062279.1 <i>S. prunicolor</i>
<i>solinT3</i>	179	Transportador MFS	86/91	WP_004003304.1 <i>S.s viridochromogenes</i>
<i>solinH</i>	327	Alfa-Beta Hidrolase	56/69	WP_030871832.1 <i>Streptomyces</i> sp. NRRL S-37
<i>solinJ</i>	246	Alfa-Beta Hidrolase (3-oxoadipate enol-lactonase)	83/91	WP_004003305.1 <i>S. viridochromogenes</i>
<i>solinK</i>	45	Aspartato 1-decarboxilase	85/90	WP_004003306.1 <i>S. viridochromogenes</i>
<i>solinS</i>	374	MbtH proteína (Domínio presente em NRPS)	89/92	WP_004003307.1 <i>S. viridochromogenes</i>
<i>solinA</i>	308	NRPS Sintase (Domínio Condensação + Adenilação + Fosfopanteteína PCP)	75/81	WP_004003308.1 <i>S. viridochromogenes</i>
<i>solinB</i>	253	NRPS Sintase (Domínio Condensação + Adenilação + Fosfopanteteína PCP + Epimerização)	79/84	WP_016433565.1 <i>Streptomyces</i> sp. HGB0020
<i>solinC</i>	427	NRPS Sintase (Domínio Condensação + Adenilação + Fosfopanteteína PCP + Epimerização)	99/99	WP_031120697.1 <i>Streptomyces</i> sp. NRRL WC-3641
<i>solinD</i>	312	NRPS Sintase (Domínio Condensação + Adenilação + Fosfopanteteína PCP)	84/88	WP_004003311.1 <i>S. viridochromogenes</i>
<i>solinT4</i>	288	Transportador tipo ABC	71/86	WP_030614041.1 <i>S. fulvoviolaceus</i>
<i>solinL</i>	204	Formiltransferase (Metionina síntese)	91/95	WP_004003313.1 <i>S. viridochromogenes</i>
<i>solinM</i>	150	Lisina – Ornitina Monooxygenase	85/89	WP_030048561.1 <i>S. peruviansis</i>
<i>solinQ</i>	233	Metiltransferase SAM-dependente	74/84	WP_030648609.1 <i>S. rimosus</i>
<i>solin1</i>	357	ATPase transportadora de metais (Carbonato desidratase)	81/86	WP_023587981.1 <i>S. thermolilacinus</i>
<i>solin2</i>	353	ATPase (Transportadora de Íons)	30/44	WP_003930948.1 <i>Mycobacterium vaccae</i>
<i>solin3</i>	544	Proteína hipotética secretada (Domínio associada a metais pesados)	61/71	WP_016325312.1 <i>S. lividans</i>
<i>solinG1</i>	447	Glicoamilase (Glicohidrolase)	93/96	WP_031110310.1 <i>Streptomyces</i> sp. NRRL S-146
<i>solinG2</i>	429	Regulador repressor no metabolismo de açúcares (Domínio DeoR)	64/74	WP_007028145.1 <i>Amycolatopsis decaplanina</i>
<i>solinG3</i>	113	Glicosamine-6-fosfato deaminase (isomerase)	24/34	WP_018720630.1 <i>Salinispora pacifica</i>
<i>solinG4</i>	155	Tagatose – fructose fosfatoaldolase	89/92	WP_031139431.1 <i>S. flavovariabilis</i>
<i>solin4</i>	143	50S rRNA Metiltransferase	99/99	WP_031120712.1 <i>Streptomyces</i> sp. NRRL WC-3641
<i>solinN</i>	72	Proteína de reparo de DNA (Oxygenase)	89/91	WP_003988278.1 <i>S. viridochromogenes</i>
<i>solinR1</i>	1092	Regulador repressor ORF quinase (ROK)	98/99	WP_029181406.1 <i>S. chartreusis</i>
<i>solinR2</i>	1519	Regulador transcricional LysR	91/93	WP_010043268.1 <i>S. chartreusis</i>

4.5.8 Cluster do aminociclitol

Os aminociclitolis são um grupo amplo de produtos naturais, com diversa atividade biológica de importância na indústria clínica e agrícola. O *cluster* do antibiótico da família dos aminociclitolis encontrado em *Streptomyces olindensis* (Figura 26) contém seis genes potenciais para biossíntese do anel de aminociclitol. Entre estes genes encontram-se as *ORF's amnA* e *amnR* que codificam para uma acetilase de N-acetilglucosamina e uma aminotransferase respectivamente, as quais cumpririam a função já descrita dentro do grupo dos aminoglicosídeos, do seu possível homólogo o gene *stsC* que codifica para uma glutamina aminotransferase encontrada na via biossintética da estreptomicina (AHLERT et al., 1997).

Adicionalmente, para a síntese do núcleo da molécula, aminociclitol, a via contém quatro principais genes, sendo um deles (*amnE*) homólogo ao gene *btrC*, encontrado em *Bacillus circulans*, que codifica para o 2-deoxy-scylo-inosose, enzima crucial para os primeiros passos de formação do anel. Dentro destes quatro genes também se encontram o *amnQ*, *amnN* e *amnM* que formam o aminociclitol e ajudam na óxido-redução e desidrogenação do mesmo, sendo encontrados na maioria dos *clusters* de produção de aminoglicosídeos (CHOI; WU; CHOENG, 2008). Por outro lado, as *ORF's amnX*, *amnW*, *amnT* e *amnU* são transportadoras da família MFS ou ABC que, mediante mecanismos de efluxo, conferem resistência à bactéria contra o composto que está sendo produzido sob a expressão desses genes. Além disso, também possui um gene regulador tipo SARP (*amnZ*) como potencial ativador da via metabólica (Tabela 14).

Figura 26 Organização do *cluster* gênico envolvido na biossíntese de uma molécula da família dos aminociclitolis

Tabela 14 ORFs preditas no *cluster* biossintético de aminociclitol

ORF	No. a.a.	Função Proposta	% Ident./ Simil.	Homólogo e Origem
<i>orf1</i>	458	Treonina desidratase	98/99	WP_010045676 <i>Streptomyces chartreusis</i>
<i>amnX</i>	325	Transportador tipo ABC	95/97	WP_010045678 <i>Streptomyces chartreusis</i>
<i>amnW</i>	249	Permease/Transportador tipo ABC	83/90	WP_003954881 <i>Streptomyces clavuligerus</i>
<i>orf2</i>	165	Fator de alongação na Transcrição GreA	93/96	WP_006132146 <i>Streptomyces gancidicus</i>
<i>amnV</i>	132	Proteína de membrana	76/85	WP_007384366 <i>Streptomyces sviveus</i>
<i>amnA</i>	286	N-acetilglucosamina-PI de-N-acetilase	99/99	WP_010045685 <i>Streptomyces chartreusis</i>
<i>orf3</i>	89	Proteína Hipotética transmembranar	81/86	WP_010045686 <i>Streptomyces chartreusis</i>
<i>amnZ</i>	1067	Proteína regulatória Tipo SARP	94/96	WP_010045687 <i>Streptomyces chartreusis</i>
<i>orf4</i>	176	Proteína Hipotética	83/92	WP_007385526 <i>Streptomyces sviveus</i>
<i>amnS</i>	303	Álcool desidrogenase GroES-like	83/88	WP_007385525 <i>Streptomyces sviveus</i>
<i>orf5</i>	159	Proteína Hipotética	88/92	WP_006603460 <i>Streptomyces auratus</i>
<i>amnQ</i>	411	2-epi-5-epi-valiolone sintase Aminociclitol (SalQ)	83/90	WP_006603459 <i>Streptomyces auratus</i>
<i>amnN</i>	249	Ciclitol oxidoreductase (SalN)	82/89	WP_007385522 <i>Streptomyces sviveus</i>
<i>amnE</i>	227	2-deoxyglycose-6-fosfatase (sintase)	82/87	WP_006603457 <i>Streptomyces auratus</i>
<i>amnB</i>	346	Glicose quinase (Família ROK)	18/28	WP_023537703 <i>Streptomyces niveus</i>
<i>amnM</i>	362	Ciclitol desidrogenase	80/87	WP_006603455 <i>Streptomyces auratus</i>
<i>amnO</i>	276	Epimerase (Xilose Isomerase)	83/86	WP_007385519 <i>Streptomyces sviveus</i>
<i>orf6</i>	345	Álcool desidrogenase (Treonina desidrogenase)	86/89	WP_006603453 <i>Streptomyces auratus</i>
<i>amnT</i>	416	Transportador MFS (Major Facilitator Superfamily)	15/28	WP_020866091 <i>Streptomyces rapamycinicus</i>
<i>orf7</i>	258	Esterase (3-oxoadipate enol-lactona hidrolase)	89/92	WP_007385516 <i>Streptomyces sviveus</i>
<i>amnR</i>	425	Aciltransferase - Aminotransferase	81/85	WP_007385515 <i>Streptomyces sviveus</i>
<i>amnU</i>	398	Transportador MFS (Major Facilitator Superfamily)	84/89	WP_007385514 <i>Streptomyces sviveus</i>
<i>amn8</i>	404	Aciltransferase - Aminotransferase	86/88	WP_007385513 <i>Streptomyces sviveus</i>
<i>amn9</i>	147	Proteína Hipotética	29/39	KOX11920 <i>Saccharothrix</i> sp. NRRL B-16348
<i>amnK</i>	63	Xilulose quinase	37/50	<i>Acetobacterium bakii</i> KNZ41063
<i>amnJ</i>	677	N-acyl glicosamina 2 epimerase	84/90	WP_019055850 <i>S. prunicolor</i>
<i>amnI</i>	425	Glicosiltransferase	97/97	WP_010045693 <i>S. chartreusis</i>
<i>amnH</i>	549	Glicosiltransferase	96/97	WP_020272420 <i>S. afghaniensis</i>
<i>amnG</i>	397	Dolcitol-fosfato mannosiltransferase	96/97	WP_020272419 <i>S. afghaniensis</i>
<i>amn10</i>	803	Proteína hipotética	94/97	WP_003992624 <i>S. viridochromogenes</i>

4.5.9 Policetídeos preditos em *S. olindensis*

Os policetídeos são uns dos grupos mais explorados como produtos naturais pela diversidade estrutural e biológica. Em *S. olindensis* foram descritos cinco clusters policetídeos, sendo três do tipo PKS I e dois do tipo PKS II, incluindo a cosmomicina D. Além, foram encontrados quatro clusters híbrido que contêm ORFs envolvidas na síntese de policetídeos. Um destes clusters apresentou homologia com um poliéter (Tetronomicina), apresentando características típicas dos PKS I como enzimas multimodulares, e a presença das enzimas epóxido hidrolase e epoxidase encarregadas da ciclização da molécula policetídic final (DEMYDCHUK et al., 2008).

4.5.9.1 Anotação do cluster biossintético do Policetídeo Tipo I – Poliéter-like

Dentro das PKS tipo I encontra-se um grupo de produtos chamados poliéteres ionóforos que transportam íons através da membrana causando a despolarização e morte das células alvo, sendo de alta importância na área pecuária e como antiparasitários especialmente por sua atividade contra a malária (LEADLAY et al., 2001).

Os policetídeos tipo I da classe dos poliéteres ionóforos, são compostos que têm mostrado uma alta atividade antiparasitária e ação de amplo espectro antifúngica, antiviral, herbicida, anti-inflamatória e imunorreguladora. O modo de ação destas moléculas é o transporte de íons através das membranas celulares o que leva à despolarização das membranas e posterior morte celular (WANG et al., 2011). Mais de 100 poliéteres têm sido descritos na literatura, mas até hoje permanecem algumas questões sobre a biossíntese das moléculas, o que representa um escopo para contribuir com o entendimento da via metabólica deste tipo de policetídeos. Neste caso, o software antiSMASH utilizando o *ClusterBlast* integrado, determinou uma alta similaridade do *cluster* biossintético, com a via de produção da tetronomicina, outro poliéter caracterizado extensivamente, assim que a descrição deste *cluster* foi comparada com esta via metabólica já descrita.

No caso do potencial *cluster* do policetídeo do tipo poliéter foram descritas 42 possíveis *ORFs* (Figura 27), entre estes, três genes transportadores potencialmente responsáveis pela resistência ao composto (*polT1*, *polT2* e *polT3*). Para a regulação

ausência de uma enzima para a produção dos núcleos assimétricos da molécula indica uma diferença em sua estrutura final, além da presença no *cluster* 23 de um módulo ACP no final de cada enzima PKS, responsável pela liberação da cadeia policetíica após a extensão da molécula (DEMYDCHUK et al., 2008) o que comumente ocorre nas vias de biossíntese de poliéteres. Segundo o software antiSMASH, os monômeros que integram a molécula são inicialmente malonil-coA, exceto por três inserções de metilmalonil-coA nos módulos codificados pelos genes *polC* e *polE*. Estes dados serviram para a elucidação da estrutura da molécula em análise posterior.

As *ORF polY1* e *polY2* foram identificadas como uma epóxido hidrolase e uma epoxidase respectivamente, genes frequentemente presentes nas vias biossintéticas de poliéteres ionóforos, como as *ORFs tmnB* e *tmnC* da tetromicina; sendo muito utilizados para a identificação desta classe de PKS tipo I. Adicionalmente, a via biossintética descrita no genoma de *S. olindensis* possui vários genes para modificações pós PKS dentro da molécula, como metiltransferases, glicosiltransferases e acetiltransferases que devem ser analisados para determinar a função específica destas enzimas. Por último, a presença de proteínas hipotéticas desconhecidas deve ser estudada para determinar se são essenciais na síntese ou não do antibiótico poliéter.

Tabela 15 Funções preditas do *cluster* biossintético para a produção do Policetídeo Tipo I – Poliéter-like

<i>ORF</i>	a.a.	Função Proposta	%Iden/ Simil.	Homólogo e Origem
<i>polI</i>	500	Desidrogenase (Oxidoreductase)	96/99	WP_003994524.1 <i>S. viridochromogenes</i>
<i>polJ</i>	447	Aciltransferase	89/94	WP_003994523.1 <i>S. viridochromogenes</i>
<i>polG1</i>	872	β -galactosidase (Glicosiltransferase)	92/95	WP_020275801.1 <i>S. afghaniensis</i>
<i>polG2</i>	406	Glicose-1-fosfato adenilyltransferase	97/98	WP_029180985.1 <i>S. chartreusis</i>
<i>polG3</i>	383	Glicosiltransferase Grupo 1	96/98	WP_010037338.1 <i>S. chartreusis</i>
<i>pol2</i>	245	Proteína Hipotética (Domínio Ferro Ferrico redutase)	86/89	WP_029180986.1 <i>S. chartreusis</i>
<i>polT1</i>	414	Proteína de membrana (Transportadora de metabólitos DMT)	72/75	WP_010037343.1 <i>S. chartreusis</i>
<i>polG4</i>	424	Transglicosilase (Peptidase)	86/92	WP_020270077.1 <i>S. afghaniensis</i>
<i>polG5</i>	50	6-phosphogluconato desidrogenase	92/95	WP_019068227.1 <i>Streptomyces</i> sp. R1-NS-10
<i>polG6</i>	479	Proteína Hipotética (Domínio	20/21	WP_014054718.1

		Galacturonosiltransferase)		<i>S. violaceusniger</i>
polK	114	Acetiltransferase	89/94	WP_003994515.1 <i>S. viridochromogenes</i>
pol3	139	Aspartato α -decarboxilase	96/97	WP_014676382.1 <i>S. hygrosopicus</i>
pol4	75	Proteína Hipotética	44/51	WP_006350895.1 <i>S. tsukubaensis</i>
polR1	888	Regulador transcricional hipotético Família Lux R	61/73	BAE93719.1 <i>Streptomyces</i> sp. NRRL 11266
polX	263	Tioesterase (Metabólitos secundários)	78/84	BAE93720.1 <i>Streptomyces</i> sp. NRRL 11266
pol5	75	Peptídeo Sintase	25/38	WP_013002070.1 <i>S. scabiei</i>
polL	278	2 oxoacid-desidrogenase aciltransferase	75/82	BAE93721.1 <i>Streptomyces</i> sp. NRRL 11266
polA	4968	Policetídeo sintase (KS-AT-ACP-KS-AT-DH-KR-ACP-KS-AT-DH-ER-KR-ACP)	49/59	CAQ64686.1 <i>S. lasaliensis</i>
pol6	223	Proteína Hipotética de membrana	79/89	WP_018844985.1 <i>Streptomyces</i> sp. CNS335
pol7	173	Proteína Hipotética Ribossomal	68/77	BAF73716.1 <i>Streptomyces</i> sp. NRRL 11266
polM	500	Proteína Hipotética (Monooxigenase)	82/89	WP_027774545.1 <i>Streptomyces</i> sp. CNQ329
polB	1582	Policetídeo sintase (KS-AT-KR-ACP)	44/55	WP_014057310.1 <i>S. violaceusniger</i>
polY1	146	Epóxido hidrolase/ciclase	76/84	WP_027774440.1 <i>Streptomyces</i> sp. CNQ329
polN	301	Metiltransferase	85/93	BAE93727.1 <i>Streptomyces</i> sp. NRRL 11266
polC	1664	Policetídeo Sintase (KS-AT-KR-ACP)	53/63	P_020874117.1 <i>S. rapamycinicus</i>
polD	3666	Policetídeo Sintase (KS-AT-DH-KR-ACP-KS-AT-KR-ACP)	42/53	ADY00168.1 <i>S. autolyticus</i>
polE	3828	Policetídeo Sintase (KS-AT-DH-ER-KR-ACP-KS-AT-DH-KR-ACP)	69/78	BAE93731.1 <i>Streptomyces</i> sp. NRRL 11266
polY2	463	Epoxidase	86/90	BAE93732.1 <i>Streptomyces</i> sp. NRRL 11266
polO	78	Ferredoxina	44/56	WP_026252613.1 <i>Streptomyces</i> sp. PsTaAH-124
polP	42	Citocromo P450	91/94	WP_027757645.1 <i>Streptomyces</i> sp. CNH099
polQ	400	3-oxoacil-ACP sintase	27/43	WP_003973133.1 <i>S. lividans</i>
polR2	345	MarR Regulador Transcricional	58/74	ACR50786.1 <i>S. longisporoflavus</i>
polU	54	Metoximalonil-ACP-sintase	85/91	WP_027757643.1 <i>Streptomyces</i> sp. CNH099
polV	194	Aciltransferase	75/82	WP_027770069.1 <i>Streptomyces</i> sp. CNQ329
polT2	632	Transportador Tipo 2-ABC	68/82	ACR50787.1 <i>S. longisporoflavus</i>
polT3	345	Transportador Tipo ABC (ATP Binding)	78/87	ACR50788.1 <i>S. longisporoflavus</i>
polS	543	Proteína Regulatória Tipo SARP (Domínio BTAD)	88/93	WP_027757641.1 <i>Streptomyces</i> sp. CNH099
polF	296	Policetídeo Sintase (KS-DH-ACP-TE)	69/76	WP_027770067.1 <i>Streptomyces</i> sp. CNQ329
polH	230	Policetídeo Sintase (AT-ACP-KS-TRANSAT-ACP-KS-TRANSAT-ACP-KR)	73/78	WP_020610477.1 <i>Streptomyces</i> sp. CNY243
polW	755	Acil-coA AMP – Ligase	82/89	WP_027770065.1 <i>Streptomyces</i> sp. CNQ329
polR3	2466	TetR Regulador Transcricional	78/82	WP_027745625.1 <i>Streptomyces</i> sp. CNT371
pol8	582	Proteína Hipotética de membrana (Domínio YhgE/Pip)	72/81	WP_027770063.1 <i>Streptomyces</i> sp. CNQ329

4.6 Amplificação dos genes SARP

Além dos genes reguladores tipo SARP encontrados nas novas vias biossintéticas, no *cluster* de produção de cosmomicina D, foi identificado um gene regulador pertencente à mesma família de proteínas reguladoras com o domínio BTAD conservado. Este gene também foi amplificado para assim, ter um controle na superexpressão das vias com um produto natural já conhecido e descrito. As condições da amplificação foram padronizadas realizando uma gradiente de temperatura estabelecendo 60°C como valor ótimo para a obtenção do amplicon.

Figura 28 A). Padronização das condições de amplificação utilizando uma gradiente de temperatura. (60°C - 72°C) (SARP Cos: Gene Regulador na via da Cosmomicina D – 960 pb; SARP Pol: Gene Regulador na via do Poliéter-like – 948 pb). B). Amplificação do fragmento do gene regulador SARP do *cluster* do aminoglicosídeo. SARP Amin: 3377 pb).

Após a qualificação, foi sugerido deter o trabalho sob super-expressão e decidiu-se concentrar sob o trabalho *in-silico* no genoma de *S. olindensis*. As atividades realizadas até o momento permitiram identificar diversas vias biossintéticas de interesse biotecnológico dentro de um único micro-organismo, revelando o potencial que há nos genomas de actinobactérias e a necessidade de ferramentas interdisciplinares para a descrição de novas rotas metabólicas.

5 CONCLUSÕES

- A descrição das características físicas permitiu comparar a topologia do cromossomo de *S. olindensis* com genomas de referência, atribuindo à linearidade *in-silico* e identificando proteínas envolvidas na plasticidade e receptividade do cromossomo, encontrando a possível inserção do *cluster* do metabólito híbrido entre melanina e nucleosídeo.
- A classificação funcional das proteínas codificadas no genoma de *Streptomyces olindensis* DAUFPE 5622, atribui mais do 50% das regiões codificadoras à funções do metabolismo primário e 2.8% das proteínas foram relacionadas com metabolismo secundário.
- A análise de genômica comparativa, utilizando a linhagem *S. olindensis* DAUFPE 5622 e quatro espécies próximas altamente conhecidas, permitiu elucidar a conformação do core genoma em bactérias deste gênero, relatando vias de metabolismo primário, e alguns metabólitos secundários amplamente distribuídos em actinobactérias como sideróforos e exopolissacarídeos.
- Utilizando os grupos de ortólogos da base de dado do KEGG, conseguiu-se determinar as funções dos elementos únicos presentes em *S. olindensis*, destacando-se as proteínas envolvidas na produção de metabólitos secundários como policetídeo sintases, NRP sintases, terpeno sintases entre outros.
- A ferramenta antiSMASH permitiu a identificação da maioria dos *clusters* biossintéticos, facilitando a caracterização das vias metabólicas visando a produção de novos produtos naturais bioativos.
- Realizando a anotação funcional e caracterização dos *clusters* metabólicos, logrou-se realizar uma comparação com vias metabólicas já conhecidas para assim, evitar o redescobrimento de moléculas descritas previamente.

- A identificação e anotação das 34 vias biossintéticas permitiu determinar a presença de genes reguladores, selecionando os clusters biossintéticos a serem ativados.

REFERÊNCIAS*

- AHLERT, J. et al. Identification of stsC, the gene encoding the L-glutamine:scyllo-inosose aminotransferase from streptomycin-producing *Streptomyces*. **Archives of Microbiology**, v. 168, n. 2, p. 102–113, 1997.
- AIGLE, B. et al. Genome mining of *Streptomyces ambofaciens*. **Journal of Industrial Microbiology & Biotechnology**, v. 41, n. 2, p. 251–263, 2014.
- AIGLE, B.; CORRE, C. **Waking up *Streptomyces* secondary metabolism by constitutive expression of activators or genetic disruption of repressors**. [s.l.] Elsevier Inc., 2012. v. 517
- AJITHKUMAR, V.; PRASAD, R. The activator/repressor protein DnrO of *Streptomyces peucetius* binds to DNA without changing its topology. **International Journal of Biological Macromolecules**, v. 46, n. 3, p. 380–384, 2010.
- ALAM, M. T. et al. Genome-based phylogenetic analysis of *Streptomyces* and its relatives. **Molecular Phylogenetics and Evolution**, v. 54, n. 3, p. 763–772, 2010.
- ALAM, M. T. et al. Comparative genome-scale metabolic modeling of actinomycetes: The topology of essential core metabolism. **FEBS Letters**, v. 585, n. 14, p. 2389–2394, 2011.
- ARNISON, P. G. et al. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. **Natural Product Reports**, v. 30, n. 1, p. 108–160, 2013.
- AZIZ, R. K. et al. The RAST Server: rapid annotations using subsystems technology. **BMC Genomics**, v. 9, n. 1, p. 75, 2008.
- BALTZ, R. H. *Streptomyces* and *Saccharopolyspora* hosts for heterologous expression of secondary metabolite gene clusters. **Journal of Industrial Microbiology & Biotechnology**, v. 37, n. 8, p. 759–772, 2010.
- BARANASIC, D. et al. Draft genome sequence of *Streptomyces rapamycinicus* strain NRRL5491, the producer of the immunosuppressant rapamycin. **Genome Announcements**, v. 1, n. 4, p. 8766, 2013.
- BENTLEY, S. et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). **Nature**, v. 417, n. 6885, p. 141–147, 2002.
- BÉRDY, J. Bioactive microbial metabolites. **The Journal of Antibiotics**, v. 58, n. 1, p. 1–26, 2005.

* De acordo com:
ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023**: informação e documentação: referências: elaboração. Rio de Janeiro, 2002.

BÉRDY, J. Thoughts and facts about antibiotics: Where we are now and where we are heading. **The Journal of Antibiotics**, v. 65, n. 8, p. 441–441, 2012.

BIBB, M. J. Regulation of secondary metabolism in streptomycetes. **Current Opinion in Microbiology**, v. 8, n. 2, p. 208–215, 2005.

BIEBER, L. W. et al. The anthracyclenic complex retamycin,. Structure. Determination of the major constituents. **Journal of Natural Products**, v. 2, p. 385–388, 1989.

BLAŽIČ, M. et al. Annotation of the modular polyketide synthase and nonribosomal peptide synthetase gene clusters in the genome of *Streptomyces tsukubaensis* NRRL18488. **Applied and Environmental Microbiology**, v. 78, n. 23, p. 8183–8190, 2012.

BODE, H. B. et al. Big effects from small changes : possible ways to explore nature ' s chemical diversity. **ChemBioChem**, v. 3, p. 619–627, 2002.

BORISOVA, S. A. et al. Biosynthesis of rhizocticins, antifungal phosphonate oligopeptides produced by *Bacillus subtilis* ATCC6633. **Chemistry & Biology**, v. 17, n. 1, p. 28–37, 2010.

BURSY, J. et al. Synthesis and uptake of the compatible solutes ectoine and 5-hydroxyectoine by *Streptomyces coelicolor* A3(2) in Response to Salt and Heat Stresses. **Applied and Environmental Microbiology**, v. 74, n. 23, p. 7286–7296, 2008.

CANE, D. E. E.; IKEDA, H. Exploration and mining of the bacterial terpenome. **Accounts of Chemical Research**, v. 45, n. 3, p. 463–472, 2012.

CHALLIS, G. L. Exploitation of the *Streptomyces coelicolor* A3(2) genome sequence for discovery of new natural products and biosynthetic pathways. **Journal of Industrial Microbiology & Biotechnology**, v. 41, n. 2, p. 219–232, 2014.

CHANDRA, G.; CHATER, K. F. Developmental biology of *Streptomyces* from the perspective of 100 actinobacterial genome sequences. **FEMS Microbiology Reviews**, v. 38, n. 3, p. 345–379, 2014.

CHEN, D. et al. Improvement of FK506 Production in *Streptomyces tsukubaensis* by genetic enhancement of the supply of unusual polyketide extender units via utilization of two distinct site-specific recombination systems. **Applied and Environmental Microbiology**, v. 78, n. 15, p. 5093–5103, 2012.

CHOI, W. S.; WU, X.; CHOENG, Y. Genetic organization of the putative salbostatin biosynthetic gene cluster including the 2- epi -5- epi -valiolone synthase gene in *Streptomyces albus* ATCC 21838. **Applied Genetics and Molecular Biotechnology**, p. 637–645, 2008.

CIMERMANCIC, P. et al. Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. **Cell**, v. 158, n. 2, p. 412–421, 2014.

CIONI, J. P. et al. Cyanohydrin phosphonate natural product from *Streptomyces regensis*. **Journal of Natural Products**, v. 77, n. 2, p. 243–249, 2014.

COLE, S. P. et al. Biosynthesis of the antibiotic actinorhodin. Analysis of blocked mutants of *Streptomyces coelicolor*. **Journal of Antibiotics (Tokyo)**, v. 40, n. 3, p. 340–347, 1987.

CONTRERAS HERNÁNDEZ C. A. **Caracterização de genes biossintéticos do antitumoral cosmomicina D**. 2013. 105 f. Dissertação (Mestrado em Microbiologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2013.

DELCHER, A. L. et al. Identifying bacterial genes and endosymbiont DNA with Glimmer. **Bioinformatics**, v. 23, n. 6, p. 673–679, 2007.

DEMAIN, A. L. Importance of microbial natural products and the need to revitalize their discovery. **Journal of Industrial Microbiology & Biotechnology**, v. 41, n. 2, p. 185–201, 2014.

DEMAIN, A. L.; ADRIO, J. L. Contributions of Microorganisms to Industrial Biology. **Molecular Biotechnology**, v. 38, n. 1, p. 41–55, 2008.

DEMYDCHUK, Y. et al. Analysis of the tetronomycin gene cluster: insights into the biosynthesis of a polyether tetronate antibiotic. **ChemBioChem**, v. 9, n. 7, p. 1136–1145, 2008.

DOROGHAZI, J. R.; BUCKLEY, D. H. Intraspecies comparison of *Streptomyces pratensis* genomes reveals high levels of recombination and gene conservation between strains of disparate geographic origin. **BMC genomics**, v. 15, n. 1, p. 970, 2014.

DOROGHAZI, J. R.; METCALF, W. W. Comparative genomics of actinomycetes with a focus on natural product biosynthetic genes. **BMC Genomics**, v. 14, n. 1, p. 611, 2013.

DUNN, B. J.; KHOSLA, C. Engineering the acyltransferase substrate specificity of assembly line polyketide synthases. **Journal of the Royal Society, Interface / the Royal Society**, v. 10, n. 85, p. 297, 2013.

DURHAM, A. M. et al. EGene: a configurable pipeline generation system for automated sequence analysis. **Bioinformatics**, v. 21, n. 12, p. 2812–2813, 2005.

EDDY, S. R. Accelerated profile HMM searches. **PLoS Computational Biology**, v. 7, n. 10, p. e1002195, 2011.

EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acid Research**, v. 32, n. 5, p. 1792–1797, 2004.

ENGEL, R. Phosphonates as analogues of natural phosphates. **Chemical Reviews**, v. 77, p. 349–367, 1977.

EVANS, B. S. et al. Discovery of the antibiotic phosacetamycin via a new mass spectrometry-based method for phosphonic acid detection. **ACS chemical biology**, v. 8, n. 5, p. 908–913, 2013.

FELNAGLE, E. A. et al. Nonribosomal peptide synthetases involved in the production of medically relevant natural products. **Molecular Pharmaceutics**, v. 5, n. 2, p. 191–211, 2008.

FINKING, R.; MARAHIEL, M. A. Biosynthesis of Nonribosomal Peptides. **Annual Review of Microbiology**, v. 58, n. 1, p. 453–488, 2004.

FINN, R. D.; CLEMENTS, J.; EDDY, S. R. HMMER web server: interactive sequence similarity searching. **Nucleic Acids Research**, v. 39, p. W29–37, 2011.

FISCHBACH, M. A; WALSH, C. T. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. **Chemical Reviews**, v. 5, p. 3468–3496, 2006.

FUNA, N. et al. Biosynthesis of hexahydroxyperylenequinone melanin via oxidative aryl coupling by cytochrome P-450 in *Streptomyces griseus*. **Journal of Bacteriology**, v. 187, n. 23, p. 8149–8155, 2005.

FURLAN, R. L. A et al. DNA-binding properties of cosmomycin D, an anthracycline with two trisaccharide chains. **The Journal of Antibiotics**, v. 57, n. 10, p. 647–54, 2004.

GAN, H. et al. Comparative genomic analysis of six bacteria belonging to the genus *Novosphingobium*: insights into marine adaptation, cell-cell signaling and bioremediation. **BMC Genomics**, v. 14, n. 1, p. 431, 2013.

GARRIDO, L. M. et al. Insights in the glycosylation steps during biosynthesis of the antitumor anthracycline cosmomycin: characterization of two glycosyltransferase genes. **Applied Microbiology and Biotechnology**, v. 73, n. 1, p. 122–131, 2006.

GETSIN, I. et al. Comparative genomics of transport proteins in developmental bacteria: *Myxococcus xanthus* and *Streptomyces coelicolor*. **BMC Microbiology**, v. 13, n. 1, p. 279, 2013.

GOMEZ-ESCRIBANO, J. P.; BIBB, M. J. Heterologous expression of natural product biosynthetic gene clusters in *Streptomyces coelicolor*: from genome mining to manipulation of biosynthetic pathways. **Journal of Industrial Microbiology & Biotechnology**, v. 41, n. 2, p. 425–431, 2014.

GOODFELLOW, M.; WILLIAMS, S. T. Ecology of actinomycetes. **Annual Review of Microbiology**, v. 37, n. 41, p. 189–216, 1983.

GUO, J. et al. High-level production of melanin by a novel isolate of *Streptomyces kathirae*. **FEMS Microbiology Letters**, v. 357, n. 1, p. 85–91, 2014.

GUYET, A. et al. Regulation of the clpP1-clpP2 operon by the pleiotropic regulator AdpA in *Streptomyces lividans*. **Archives of Microbiology**, v. 195, n. 12, p. 831–841, 2013.

HARVEY, A. L.; EDRADA-EBEL, R.; QUINN, R. J. The re-emergence of natural products for drug discovery in the genomics era. **Nature Reviews Drug Discovery**, v. 14, n. 2, p. 111–129, 2015.

HERTWECK, C. The biosynthetic logic of polyketide diversity. **Angewandte Chemie International edition**, v. 48, n. 26, p. 4688–4716, jan. 2009.

HOPWOOD, D. A. Forty years of genetics with *Streptomyces*: from in vivo through in vitro to in silico. **Microbiology**, v. 145, p. 2183–2202, 1999.

HOPWOOD, D. A. Soil To Genomics: The *Streptomyces* Chromosome. **Annual Review of Genetics**, v. 40, p. 1–23, 2006.

HUANG, C. H. et al. The telomeres of *Streptomyces* chromosomes contain conserved palindromic sequences with potential to form complex secondary structures. **Molecular Microbiology**, v. 28, n. 5, p. 905–916, 1998.

HUTCHINGS, M. I. et al. The E cell envelope stress response of *Streptomyces coelicolor* is influenced by a novel lipoprotein, CseA. **Journal of Bacteriology**, v. 188, n. 20, p. 7222–7229, 2006.

HWANG, K.-S. et al. Systems biology and biotechnology of *Streptomyces* species for the production of secondary metabolites. **Biotechnology Advances**, v. 32, n. 2, p. 255–268, 2014.

IKEDA, H. et al. Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. **Nature Biotechnology**, v. 21, n. 5, p. 526–531, 2003.

JANDA, J. M.; ABBOTT, S. L. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. **Journal of Clinical Microbiology**, v. 45, n. 9, p. 2761–2764, 2007.

JIANG, J.; HE, X.; CANE, D. E. Geosmin biosynthesis. *Streptomyces coelicolor* germacradienol/germacrene D synthase converts farnesyl diphosphate to geosmin. **Journal of the American Chemical Society**, v. 128, n. 25, p. 8128–8129, 2006.

JOHNSON, M. et al. NCBI BLAST: a better web interface. **Nucleic Acids Research**, v. 36, p. W5–W9, 2008.

JU, K.-S.; DOROGHAZI, J. R.; METCALF, W. W. Genomics-enabled discovery of phosphonate natural products and their biosynthetic pathways. **Journal of Industrial Microbiology & Biotechnology**, v. 41, n. 2, p. 345–356, 2014.

KÄLL, L.; KROGH, A.; SONNHAMMER, E. L. A combined transmembrane topology and signal peptide prediction method. **Journal of Molecular Biology**, v. 338, n. 5, p. 1027–1036, 2004.

KATO, J. et al. Biosynthesis of gamma-butyrolactone autoregulators that switch on secondary metabolism and morphological development in *Streptomyces*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 104, p. 2378–2383, 2007.

KIESER, T.; BIBB, M. J.; BUTTENER, M. J.; CHATER, K. F.; HOPWOOD, D. A. **Practical Streptomyces Genetics**. Norwich: The John Innes Foundation, 2002. p

KIM, J.-N. et al. Comparative genomics reveals the core accessory genomes of *Streptomyces* species. **Journal of Microbiology and Biotechnology**, 2015.

KOTOWSKA, M.; PAWLIK, K. Roles of type II thioesterases and their application for secondary metabolite yield improvement. **Applied Microbiology and Biotechnology**, p. 7735–7746, 2014.

KROGH, A. et al. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. **Journal of Molecular Biology**, v. 305, n. 3, p. 567–580, 2001.

KUDO, F. et al. Cloning of the pactamycin biosynthetic gene cluster and characterization of a crucial glycosyltransferase prior to a unique cyclopentane ring formation. **The Journal of Antibiotics**, v. 60, n. 8, p. 492–503, 2007.

KUDO, F.; EGUCHI, T. Biosynthetic genes for aminoglycoside antibiotics. **The Journal of Antibiotics**, v. 62, n. 9, p. 471–481, 2009.

LEADLAY, P. F. et al. Engineering of complex polyketide biosynthesis--insights from sequencing of the monensin biosynthetic gene cluster. **Journal of Industrial Microbiology & Biotechnology**, v. 27, n. 6, p. 360–367, 2001.

LEZHAVA, A et al. Physical map of the linear chromosome of *Streptomyces griseus*. **Journal of Bacteriology**, v. 177, n. 22, p. 6492–8, 1995.

LIU, G. et al. Molecular regulation of antibiotic biosynthesis in *Streptomyces*. **Microbiology and Molecular Biology Reviews**, v. 77, n. 1, p. 112–143, 2013.

MAHMUD, T. Progress in aminocyclitol biosynthesis. **Current Opinion in Chemical Biology**, v. 13, n. 2, p. 161–170, 2009.

MAHMUD, T.; LEE, S.; FLOSS, H. G. The biosynthesis of acarbose and validamycin. **Chemical record**, v. 1, n. 4, p. 300–310, 2001.

MAO, X. et al. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. **Bioinformatics**, v. 21, n. 19, p. 3787–3793, 2005.

MARTÍN, J. F.; LIRAS, P. Engineering of regulatory cascades and networks controlling antibiotic biosynthesis in *Streptomyces*. **Current Opinion in Microbiology**, v. 13, n. 3, p. 263–273, 2010.

MEDEMA, M. H. The minimum information about a biosynthetic gene cluster (MIBiG) specification. **Nature Chemical Biology**, p. 625–631, 2015.

MENCHER, R. J.; HEIM, A. H. Melanin biosynthesis by *Streptomyces lavendulae*. **Journal of Genetic Microbiology**, n. 28, p. 665–670, 1962.

METCALF, W. W.; VAN DER DONK, W. A. Biosynthesis of phosphonic and phosphinic acid natural products. **Annual Review of Biochemistry**, v. 78, p. 65–94, 2009.

MITCHELL, A. et al. The InterPro protein families database: the classification resource after 15 years. **Nucleic Acids Research**, v. 43, n. D1, p. D213–D221, 2015.

MOTAMEDI, H.; SHAFIEE, A. The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506. **European Journal of Biochemistry**, v. 256, n. 3, p. 528–534, 1998.

NAKANO, C.; HORINOUCI, S.; OHNISHI, Y. Characterization of a novel sesquiterpene cyclase involved in (+)-caryolan-1-ol biosynthesis in *Streptomyces griseus*. **Journal of Biological Chemistry**, v. 286, n. 32, p. 27980–27987, 2011.

NETT, M.; IKEDA, H.; MOORE, B. S. Genomic basis for natural product biosynthetic diversity in the actinomycetes. **Natural Product Reports**, v. 26, n. 11, p. 1362, 2009.

NEWMAN, D. J.; CRAGG, G. M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. **Journal of Natural Products**, v. 75, n. 3, p. 311–335, 2012.

NINDITA, Y. et al. The *tap-tpg* gene pair on the linear plasmid functions to maintain a linear topology of the chromosome in *Streptomyces rochei*. **Molecular Microbiology**, v. 95, n. 5, p. 846–858, 2015.

NOVAKOVA, R.; BISTAKOVA, J.; KORMANEC, J. Characterization of the polyketide spore pigment cluster whiESa in *Streptomyces aureofaciens* CCM3239. **Archives of Microbiology**, v. 182, n. 5, p. 388–395, 2004.

OCHI, K.; HOSAKA, T. New strategies for drug discovery: activation of silent or weakly expressed microbial gene clusters. **Applied Microbiology and Biotechnology**, v. 97, n. 1, p. 87–98, 2013.

OLANO, C.; MÉNDEZ, C.; SALAS, J. A. Antitumor Compounds from Marine Actinomycetes. **Marine Drugs**, v. 7, n. 2, p. 210–248, 2009.

OMURA, S. et al. Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. **Proceedings of**

the National Academy of Sciences of the United States of America, v. 98, n. 21, p. 12215–20, 2001.

OVERBEEK, R. et al. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). **Nucleic Acids Research**, v. 42, n. D1, p. D206–D214, 2014.

PAN, J.-J. et al. Biosynthesis of squalene from farnesyl diphosphate in bacteria: three steps catalyzed by three enzymes. **ACS Central Science**, v. 1, n. 2, p. 77–82, 2015.

PETERSEN, T. N. et al. SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature Methods**, v. 8, n. 10, p. 785–786, 2011.

PHELAN, R. M. et al. Engineering terpene biosynthesis in *Streptomyces* for production of the advanced biofuel precursor bisabolene. **ACS Synthetic Biology**, v. 4, n. 4, p. 393–399, 2014.

PORALLA, K.; MUTH, G.; HÄRTNER, T. Hopanoids are formed during transition from substrate to aerial hyphae in *Streptomyces coelicolor* A3(2). **FEMS Microbiology Letters**, v. 189, p. 93–95, 2000.

POWELL, S. et al. eggNOG v4.0: nested orthology inference across 3686 organisms. **Nucleic Acids Research**, v. 42, n. D1, p. D231–D239, 2014.

PRASEUTH, A P. et al. Complete sequence of biosynthetic gene cluster responsible for producing triostin A and evaluation of quinomycin-type antibiotics from *Streptomyces triostinicus*. **Biotechnol Prog**, v. 24, n. 6, p. 1226–1231, 2008.

RABYK, M. V.; OSTASH, B. O.; FEDORENKO, V. O. Gene networks regulating secondary metabolism in actinomycetes: Pleiotropic regulators. **Cytology and Genetics**, v. 48, n. 1, p. 55–67, 2014.

ROJAS, J. D. et al. Genome sequence of *Streptomyces olindensis* DAUFPE 5622, producer of the antitumoral anthracycline cosmomycin D. **Genome Announcements**, v. 2, n. 3, p. 4–5, 2014.

SADEGHI, A. et al. Diversity of the ectoines biosynthesis genes in the salt tolerant *Streptomyces* and evidence for inductive effect of ectoines on their accumulation. **Microbiological Research**, v. 169, n. 9–10, p. 699–708, 2014.

SALAS, J. A.; MÉNDEZ, C. Engineering the glycosylation of natural products in actinomycetes. **Trends in Microbiology**, v. 15, n. 5, p. 219–232, 2007.

SCHMIDT, E. W. The hidden diversity of ribosomal peptide natural products. **BMC biology**, v. 8, p. 83, 2010.

SCHMOOCK, G. et al. Functional Cross-talk between Fatty Acid Synthesis and Nonribosomal Peptide Synthesis in Quinoxaline Antibiotic-producing Streptomycetes. **Journal of Biological Chemistry**, v. 280, n. 6, p. 4339–4349, 2005.

SCHWECKE, T. et al. The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. **Proceedings of the National Academy of Sciences of the United States of America**, v. 92, n. 17, p. 7839–7843, 1995.

SCHWIENSTEK, P. et al. The complete genome sequence of the acarbose producer *Actinoplanes* sp. SE50/110. **BMC Genomics**, v. 13, n. 1, p. 112, 2012.

SEIPKE, R. F.; KALTENPOTH, M.; HUTCHINGS, M. I. *Streptomyces* as symbionts: An emerging and widespread theme? **FEMS Microbiology Reviews**, v. 36, n. 4, p. 862–876, 2012.

SIEBER, S. A; MARAHIEL, M. A. Learning from nature's drug factories: nonribosomal synthesis of macrocyclic peptides. **Journal of Bacteriology**, v. 185, n. 24, p. 7036–7043, 2003.

SIVAPERUMAL, P. et al. Melanin from marine *Streptomyces* sp. (MVCS13) with potential effect against ornamental fish pathogens of *Carassius auratus* (Linnaeus, 1758). **Biocatalysis and Agricultural Biotechnology**, v. 3, n. 4, p. 134–141, 2014.

SOUSA, T. D. S. et al. Anthracyclinones from *Micromonospora* sp. **Journal of Natural Products**, v. 75, n. 3, p. 489–493, 2012.

STARCEVIC, A. et al. Recombinatorial biosynthesis of polyketides. **Journal of Industrial Microbiology & Biotechnology**, v. 39, n. 3, p. 503–511, 2012.

TAMURA, K. et al. MEGA6: Molecular evolutionary genetics analysis version 6.0. **Molecular Biology and Evolution**, v. 30, n. 12, p. 2725–2729, 2013.

UDWARY, D. et al. Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. **Proceedings of the National Academy of Sciences** v. 104, n. 25, p. 10376–10381, 2007.

UNTERGASSER, A. et al. Primer3-new capabilities and interfaces. **Nucleic Acids Research**, v. 40, n. 15, p. 1–12, 2012.

VAN WEZEL, G. P.; MCDOWALL, K. J. The regulation of the secondary metabolism of *Streptomyces*: new links and experimental advances. **Natural Product Reports**, v. 28, n. 7, p. 1311, 2011.

VASANTHAKUMAR, A.; KATTUSAMY, K.; PRASAD, R. Regulation of daunorubicin biosynthesis in *Streptomyces peucetius* - feed forward and feedback transcriptional control. **Journal of Basic Microbiology**, v. 53, n. 8, p. 636–644, 2013.

VENTURA, M. et al. Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum. **Microbiology and Molecular Biology Reviews**, v. 71, n. 3, p. 495–548, 2007.

VERNIKOS, G. S.; PARKHILL, J. Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the *Salmonella* pathogenicity islands. **Bioinformatics**, v. 22, n. 18, p. 2196–2203, 2006.

VOLFF1, J.-N.; ALTENBUCHNER, J. MicroReview Genetic instability of the. **Molecular Microbiology**, v. 27, p. 239–246, 1998.

WALSH, C. T.; FISCHBACH, M. A. Natural products version 2.0: Connecting genes to molecules. **Journal of the American Chemical Society**, v. 132, n. 8, p. 2469–2493, 2010.

WANG, H. et al. Genetic screening strategy for rapid access to polyether ionophore producers and products in Actinomycetes. **Applied and Environmental Microbiology**, v. 77, n. 10, p. 3433–3442, 2011.

WANG, H. et al. Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes. **Proceedings of the National Academy of Sciences**, v. 111, n. 25, p. 9259–9264, 2014.

WANG, H.; VAN DER DONK, W. A. Biosynthesis of the class III lantipeptide catenulipeptin. **ACS Chemical Biology**, v. 7, p. 1529–1535, 2012.

WANG, K. C.; OHNUMA S. Isoprenyl diphosphate synthases. **Biochimica et Biophysica Acta**, v. 1529, p. 33–49, 2000.

WANG, X. et al. Characterization and analysis of an industrial strain of *Streptomyces bingchenggensis* by genome sequencing and gene microarray. **NRC research press**, v. 689, p. 677–689, 2013.

WEBER, T. et al. antiSMASH 3.0--a comprehensive resource for the genome mining of biosynthetic gene clusters. **Nucleic Acids Research**, v. 43, n. W1, p. W237–W243, 2015.

WEISSMAN, K. J. Introduction to polyketide biosynthesis. In: HOPWOOD, D.A. (Ed) **Methods in Enzymology: Complex Enzymes in Microbial Natural Product Biosynthesis, Part B: Polyketides, Aminocoumarins and Carbohydrates**. 2009, v. 459. p. 3–16.

WOLAŃSKI, M.; JAKIMOWICZ, D.; ZAKRZEWSKA-CZERWIŃSKA, J. AdpA, key regulator for morphological differentiation regulates bacterial chromosome replication. **Open Biology**, v. 2, n. 7, p. 120097, 2012.

WRIGHT, F.; BIBB, M. J. Codon usage in the G+C-rich *Streptomyces* genome. **Gene**, v. 113, n. 1, p. 55–65, 1992.

XU, L. et al. Complete genome sequence and comparative genomic analyses of the vancomycin-producing *Amycolatopsis orientalis*. **BMC Genomics**, p. 1–18, 2014.

YAMADA, Y. et al. Terpene synthases are widely distributed in bacteria. **Proceedings of the National Academy of Sciences of the United States of America**, v. 112, n. 3, p. 857–862, 2015.

YANG, C.-C. et al. Mutational Analysis of the Terminal Protein Tpg of *Streptomyces* Chromosomes: Identification of the Deoxynucleotidylation Site. **PLoS ONE**, v. 8, n. 2, p. e56322, 2013.

YU, T. W.; HOPWOOD, D. A. Ectopic expression of the *Streptomyces coelicolor* whiE genes for polyketide spore pigment synthesis and their interaction with the act genes for actinorhodin biosynthesis. **Microbiology**, v. 141, n. 11, p. 2779–2791, 1995.

YU, X. et al. Diversity and abundance of phosphonate biosynthetic genes in nature. **Proceedings of the National Academy of Sciences**, v. 110, n. 51, p. 20759–20764, 2013.

ZABURANNYI, N. et al. Insights into naturally minimised *Streptomyces albus* J1074 genome. **BMC Genomics**, v. 15, n. 1, p. 97, 2014.

ZHAO, B. et al. Biosynthesis of the Sesquiterpene Antibiotic Albaflavenone in *Streptomyces coelicolor* A3(2). **Journal of Biological Chemistry**, v. 283, n. 13, p. 8183–8189, 2008.

ZHOU, Z. et al. Genome plasticity and systems evolution in *Streptomyces*. **BMC Bioinformatics**, v. 13 Suppl 1, n. Suppl 10, p. S8, 2012.

ZOU, Z. et al. A γ -butyrolactone-sensing activator/repressor, JadR3, controls a regulatory mini-network for jadomycin biosynthesis. **Molecular Microbiology**, v. 94, n. 3, p. 490–505, 2014.

APÊNDICE A – Informação suplementar

Tabela 1A- Iniciadores usados para amplificação dos genes reguladores SARP

Identificação	Sequência 5'- 3'	Tm °C
SARP Primer Left Poliéter + XbaI	AGACA <u>TCTAGA</u> GGAAAGTACAGAAACGGAAGCA	71.9
SARP Primer Right Poliéter + EcoRI	AAATA <u>GAATTC</u> GGAGCTGACGCACCGTCT	72.4
SARP Primer Left Aminociclitol + XbaI	AAAAA <u>TCTAGAA</u> GGAACACGACCGAGACAAC	70.2
SARP Primer Right Aminociclitol + EcoRI	ATTTA <u>GAATTC</u> AGGAAGCAGGGCTGACCAC	71.7
SARP Primer Right Cosmomicina + XbaI	AATAA <u>TCTAGA</u> CCTTCATCTGCCCACACC	69.0
SARP Primer Left Cosmomicina + EcoRI	AATAA <u>GAATTC</u> CACTTTTCAGTGCAGGTGCGAA	69.9

Figura 1A - Distribuição de genes de *S. olindensis* dentro dos subsistemas do servidor RAST

APÊNDICE B - Funções preditas dos *clusters* achados em *S. olindensis*

Tabela B1 – Funções preditas do *cluster* I Butirolactona

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
<i>orfB</i>	282	Regulador transcricional XRE	71/80	<i>S. ambofaciens</i> WP_053135609
<i>afsA</i>	341	Proteína AfsA – Biossíntese fator A	90/93	<i>S. galbus</i> WP_052146399
<i>orfC</i>	218	Halo-ácido dealogenase	88/93	<i>S. galbus</i> WP_033530936
<i>orfD</i>	178	Lipoproteína	73/87	<i>S. aurantiacus</i> EPH40809
<i>orfE</i>	483	Sistema sensor-quinase	88/91	<i>S. aurantiacus</i> WP_016644256
<i>orfF</i>	232	Domínio efetor- sistema de regulação de resposta	45/55	<i>S. aurantiacus</i> WP_037663614

Tabela B2 – Funções preditas do *cluster* II Butirolactona

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
<i>orfB</i>	225	Regulador transcricional MerR	89/94	<i>S. viridochromogenes</i> WP_048587169
<i>orfC</i>	191	Proteína hipotética	84/90	<i>S. aurantiacus</i> WP_037658948
<i>orfD</i>	398	Epóxido hidrolase	88/91	<i>S. chartreusis</i> WP_010047241
<i>orfF</i>	226	Fosfatase	72/76	<i>S. iakyrus</i> WP_051814751
<i>afsA</i>	745	Proteína de biossíntese do Fator A – Proteína AfsA	18/30	<i>S. scabiei</i> WP_037693001
<i>orfF</i>	217	Regulador transcricional TetR	91/95	<i>S. iakyrus</i> WP_033310044

Tabela B3 – Funções preditas do *cluster* III Butirolactona

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
<i>orfB</i>	377	Butiril-CoA desidrogenase	26/39	<i>Rhodococcus opacus</i> WP_012689587
<i>orfC</i>	109	Sulfotransferase	96/98	<i>S. ghanaensis</i> WP_004985844
<i>orf1</i>	277	Proteína Hipotética	94/96	<i>S. iakyrus</i> WP_051814743
<i>orf2</i>	473	Proteína Hipotética	93/96	<i>S. azureus</i> GAP47891
<i>arpA</i>	245	Fator A - proteína receptora (gama butirolactona) ArpA	93/95	<i>S. afghaniensis</i> WP_020273122
<i>afsA</i>	360	Fator A – Proteína AfsA	78/83	<i>S. azureus</i> GAP47893
<i>orfF</i>	307	UDP-glicose 4-epimerase-desidratase	85/86	<i>S. azureus</i> GAP47894
<i>orfE</i>	422	Citocromo p450	92/95	<i>S. afghaniensis</i> WP_020273119
<i>orfD</i>	232	Regulador Transcricional TetR	16/24	<i>Gordonia rhizosphaera</i> WP_006339600

Tabela B4 - Funções preditas do *cluster* de geosmina

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
orfR	215	Regulador TetR	96/98	<i>S. viridochromogenes</i> WP_003993800
orf1	395	Proteína hipotética	90/93	<i>S. azureus</i> GAP48695
orfB	413	Glicerolfosforil diester fosfodiesterase	92/94	<i>S. azureus</i> GAP48696
orfA	721	Germacradienol sintase – Geosmina ciclase	82/89	<i>S. coelicolor</i> NP_630182
orfC	625	Fenoxazinona sintase	91/95	<i>S. viridochromogenes</i> WP_003993803
orfT	185	Proteína de membrana	84/86	<i>S. azureus</i> GAP48699
orfD	251	Fosfodiesterase	87/90	<i>S. azureus</i> GAP48700

Tabela B5 - Funções preditas do *cluster* de 2-Metilisoborneol

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
B	159	3-demethylubiquinone-9 3-methyltransferase	90/96	<i>S. viridochromogenes</i> WP_003987933
C	291	Fosfotransferase	75/84	<i>S. ghanaensis</i> EFE71182
1	242	Lipoproteína	86/90	<i>S. iakyrus</i> WP_033309425
A	251	2-metilisoborneol sintase (Terpeno ciclase)	77/86	<i>S. davawensis</i> WP_015662968
2	252	Hidrolase (SgaA_N_like)	75/86	<i>S. zinciresistens</i> WP_007501747

Tabela B6 - Funções preditas do *cluster* de Albaflavenona

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
1	346	Proteína hipotética	94/96	<i>S. zinciresistens</i> EGX61410
G	216	Peptidil deformilase	97/99	<i>S. viridochromogenes</i> EFL34751
A	361	Terpeno sintase	80/88	<i>S. avermitilis</i> BAC70743
B	451	Citocromo P450	87/92	<i>S. coelicolor</i> NP_629370
R	320	Regulador transcricional AraC	92/95	<i>S. viridochromogenes</i> EFL34754
C	335	Ribonucleotídeo difosfatase subunidade beta	98/99	<i>S. afghaniensis</i> EPJ37569
D	783	Ribonucleotídeo difosfatase subunidade alfa	95/97	<i>S. viridochromogenes</i> EFL34757
E	178	Acetiltransferase	84/90	<i>S. viridochromogenes</i> ELS56403
T	234	Proteína de membrana	64/77	<i>S. hygrosopicus</i> AEY91610
F	542	Acetato permease ActP	98/93	<i>S. glaucescens</i> AIS00449
U	84	Proteína de membrana	96/97	<i>S. viridochromogenes</i> EFL34760

Tabela B7 - Funções preditas do *cluster* de hopanóides

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
	215	Regulador transcricional XRE	93/94	<i>S. iakyrus</i> WP_033307737
<i>hpnO</i>	471	Aminotransferase	96/98	<i>S. azureus</i> GAP52687
<i>dxpS</i>	631	1-deoxi-D-xilulose 5-fosfate sintase (Via mevalonato)	96/98	<i>S. chartreusis</i> WP_010044657
<i>ispG</i>	385	4 hidroxilo 3 metilbut 2-en-1-il difosfato sintase	97/98	<i>S. afghaniensis</i> WP_020272221
<i>hpnH</i>	340	Proteína radical SAM	96/99	<i>S. zinciresistens</i> WP_007494130
<i>hpnG</i>	213	Fosforilase associada à síntese de hopanóides	87/91	<i>S. viridochromogenes</i> WP_048580875
<i>hpnF</i>	686	Esqualeno ciclase (hopanóides)	92/95	<i>S. viridochromogenes</i> WP_003996258
<i>ispH</i>	355	Isoprenil metilalil difosfato sintase	91/94	<i>S. avermitilis</i> WP_037645888
<i>hpnE</i>	470	Fitoeno desaturase FAD dependente	94/95	<i>S. afghaniensis</i> WP_020272215
<i>hpnD</i>	316	Esqualeno sintase	92/94	<i>S. viridochromogenes</i> WP_037885737
<i>hpnC</i>	300	Esqualeno sintase	94/96	<i>S. viridochromogenes</i> WP_003994382
	263	Transportador tipo ABC – ATP binding de ácidos teicóicos	94/98	<i>S. avermitilis</i> WP_010983094
<i>hpnM</i>	309	Transportador tipo ABC	94/96	<i>S. afghaniensis</i> WP_020272210
<i>hpnB</i>	290	Glicosiltransferase	95/97	<i>S. viridochromogenes</i> WP_048580882
	232	fosfatidiltransferase	97/98	<i>S. iakyrus</i> WP_033307723
	353	Glicerol desidrogenase	99/99	<i>S. afghaniensis</i> WP_020272207
<i>hpnI</i>	250	NDP sugar Transferase	97/98	<i>S. viridochromogenes</i> WP_003994376
	598	Transferase	93/95	<i>S. iakyrus</i> WP_033307720
<i>hpnA</i>	323	UDP-glicose 4-epimerase	93/96	<i>S. iakyrus</i> WP_033307719
	312	Transportador de cátions	85/92	<i>S. griseus</i> WP_037673386
<i>fni</i>	197	Isopentenil difosfato isomerase (IPP-DMAPP)	93/95	<i>S. viridochromogenes</i> WP_048580888
	168	Histidina – Quinase ATPase	93/98	<i>S. iakyrus</i> WP_033307716
	251	Enoil-CoA hidratase	93/95	<i>S. viridochromogenes</i> WP_003994369
	211	Glutamina Amidotransferase	71/80	<i>S. venezuelae</i> WP_015037612
	319	Regulador transcricional AraC	85/90	<i>S. leeuwenhoekii</i> WP_029381894

Tabela B8 - Funções preditas do *cluster* Terpeno 1 sem produto detectado

ORF	No. a.a.	Função Proposta	%Ident / Simil.	Homólogo e Origem No. Acesso
T	412	Proteína de membrana (MFS)	90/93	<i>S. chartreusis</i> WP_029181036
1	375	Proteína hipotética	83/83	<i>S. afghaniensis</i> EPJ37806
G	313	5-dehidro-4-deoxiglucarate desidratase	96/99	<i>S. afghaniensis</i> WP_037670495
H	269	NAD dependente desidratase-epimerase	97/98	<i>S. chartreusis</i> WP_010037916
S	258	Regulador transcricional DeoR	93/96	<i>S. viridochromogenes</i> WP_048583674
U	457	Proteína transportadora tipo ABC	95/97	<i>S. afghaniensis</i> WP_037670486
V	313	Proteína transportadora tipo ABC - ATP binding	97/98	<i>S. afghaniensis</i> WP_020274022
W	283	Proteína transportadora tipo ABC	79/88	<i>S. niveus</i> EST32559
D	331	Álcool (Treonina) desidrogenase	95/99	<i>S. viridochromogenes</i> WP_048583678
A	311	Geranilgeranil difosfato sintase / Trans-Isoprenil difosfato sintase	47/61	<i>S. albus</i> WP_040246818
B	304	Geranilgeranilglicerol fosfato / Preniltransferase	34/46	<i>Haladaptatus cibarius</i> WP_049970901
C	255	SAM-Metiltransferase	22/33	<i>Alicyclobacillus macrosporangiidus</i> WP_029423570
X	296	Proteína de resistência ao dano químico (Resposta ao estresse)	33/44	<i>Pseudomonas aeruginosa</i> WP_034022637
Y	39	Proteína induzida pelo estresse	97/100	<i>S. bicolor</i> WP_037639663
R	191	Proteína reguladora TetR	73/83	<i>S. avermitilis</i> WP_010987752

Tabela B9 - Funções preditas do *cluster* terpeno 2 sem produto detectado

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
T	148	Proteína hipotética de membrana	81/84	<i>S. sclerotialis</i> WP_051847641
B	278	Citocromo C oxidase	30/39	<i>S. niveus</i> WP_023543257
R	44	Sensor histidina quinase	69/80	<i>S. lividans</i> EOY47504
A	620	Isoprenil difosfato sintase (Esqualeno sintase)	81/87	<i>S. viridochromogenes</i> WP_003990021
1	302	Proteína hipotética	70/77	<i>S. atroolivaceus</i> WP_051709420
C	255	3 ceto acil redutase	89/94	<i>S. afghaniensis</i> WP_020272382
2	154	Proteína hipotética (Cupin)	96/97	<i>S. ambofaciensis</i> WP_053141513

Tabela B9 - Funções preditas do *cluster* terpeno 3 sem produto detectado

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
G	957	Glicosil-hidrolase	90/94	<i>S. chartreusis</i> WP_029181790
B	459	Citocromo P450	47/64	<i>S. hygrosopicus</i> WP_030829548
A	338	Sesquiterpeno ciclase (Pentaleneno sintase)	77/85	<i>S. viridochromogenes</i> WP_037891703
D	86	Proteína radical SAM metiltransferase	28/43	<i>S. ambofaciensis</i> WP_053138142
R	969	Proteína hipotética regulatória	74/81	<i>S. viridochromogenes</i> WP_039931952
E	662	Dioxigenase - Endoglicanase	91/94	<i>S. chartreusis</i> WP_010048709
T	493	Proteína transportadora MFS	92/96	<i>S. iakyrus</i> WP_033307176
U	685	Transportador ABC – ATP binding	68/74	<i>S. viridochromogenes</i> WP_004004151
S	137	Serina-Treonina proteína quinase	50/61	<i>S. rimosus</i> WP_033029166
C	467	Citocromo P450	95/97	<i>S. chartreusis</i> WP_010048715
F	368	NADH dehidrogenase (EC 1.6.99.3)	87/92	<i>S. griseus</i> WP_037679858

Tabela B10 - Funções preditas do *cluster* linaridina 1

ORF	No. a.a.	Função Proposta	%Ident./ Simil.	Homólogo e Origem No. Acesso
1	340	Cellulase	96/97	<i>S. azureus</i> GAP49060
2	407	Proteína hipotética	88/93	<i>S. afghaniensis</i> WP_020274781
3	605	Endonuclease (clivagem)	94/96	<i>S. iakyrus</i> WP_033312359
4	316	cypA Mycothiol biosynthesis (Acetiltransferase)	88/90	<i>S. viridochromogenes</i> WP_004000095
5	114	Proteína hipotética (sulfatase)	83/90	<i>S. afghaniensis</i> WP_020274785
6	274	Proteína de união ao DNA (Regulador transcricional)	91/95	<i>S. cyaneogriseus</i> WP_044382959
7	67	Proteína hipotética	78/88	<i>S. roseochromogenus</i> WP_023549164
8	56	Proteína hipotética (Regulador)	56/72	<i>Streptosporangium amethystogenes</i> WP_030917199
9	485	Sensor histidina quinase	91/95	<i>S. viridochromogenes</i> WP_003991237
10	245	Regulador do sistema de resposta (sensor histina)	94/96	<i>S. viridochromogenes</i> WP_003991236
11	58	Proteína hipotética	84/89	<i>S. chartreusis</i> WP_010036222
12	310	Protease	82/88	<i>S. aureofaciens</i> WP_052842966
13	41	Proteína hipotética	54/57	<i>Streptomyces</i> sp. NRRL F-525 WP_051801874

Tabela B11 – Funções preditas do *cluster* linaridina 2

ORF	No. a.a.	Função Proposta	%Ident./ Simil.	Homólogo e Origem No. Acesso
1	1216	5-oxoprolinase	90/94	<i>S. viridochromogenes</i> WP_053197020
2	201	ATP-binding protein	99/99	<i>Streptomyces chartreusis</i> WP_010046276
3	561	glycerol kinase	93/97	<i>Streptomyces violaceoruber</i> WP_030943290
4	265	glycerol transporter	87/94	<i>Streptomyces iakyrus</i> WP_033309741
5	548	diguanylate cyclase	93/95	<i>Streptomyces iakyrus</i> WP_051814728
6	49	hypothetical protein	89/91	<i>Streptomyces hygroscopicus</i> WP_041665018
7	396	lipid-transfer protein	89/93	<i>Streptomyces ipomoeae</i> WP_048820999
8	56	benzoylsuccinyl-CoA thiolase	91/94	<i>Streptomyces durhamensis</i> WP_037687735
9	74	hypothetical protein	28/43	<i>Bifidobacterium asteroides</i> WP_045935212
10	269	D-alanyl-D-alanine dipeptidase	89/95	<i>Streptomyces afghaniensis</i> WP_020275944
11	137	NUDIX hydrolase	91/94	<i>Streptomyces durhamensis</i> WP_031174395
12	740	glycosyl hydrolase	94/96	<i>Streptomyces chartreusis</i> WP_040904132
13	43	hypothetical protein BF14_2565	71/70	<i>Streptomyces griseus</i> KIX45660
14	655	acetoacetyl-CoA synthetase	91/94	<i>Streptomyces roseochromogenus</i> WP_023552212
15	288	hypothetical protein	88/93	<i>Streptomyces afghaniensis</i> WP_020275940
16	543	phosphoenolpyruvate-protein phosphotransferase	92/97	<i>Streptomyces griseus</i> WP_030852156
17	149	PTS glucose transporter subunit IIA	61/76	<i>Kutzneria albida</i> WP_025356243

Tabela B12 – Funções preditas do *cluster* lassopeptídeo

ORF	No. a.a.	Função Proposta	%Ident./ Simil.	Homólogo e Origem No. Acesso
1	408	Proteína hipotética	91/92	<i>S. azureus</i> GAP50213
2	658	Treonil-tRNA sintase	89/94	<i>S. griseus</i> WP_030712511
3	185	Fosforilase	99/100	<i>S. afghaniensis</i> WP_020270008
4	551	Proteína de membrana	96/98	<i>S. iakyrus</i> WP_033309632
5	769	Fator de alongação	32/50	<i>Actinoplanes missouriensis</i> WP_041830459
6	238	CDP-diacylglicerol--glicerol-3-fosfato 3-fosfatidiltransferase	95/96	<i>S. viridochromogenes</i> WP_039932077
7	302	aciltransferase	95/98	<i>S. iakyrus</i> WP_033309634
8	403	Glicosiltransferase	91/95	<i>S. ipomoeae</i> WP_009313179
9	46	Proteína hipotética	77/81	<i>S. lydicus</i> AJT67350
10	611	Asparagina sintase	76/84	<i>S. lydicus</i>

				WP_046928238
11	83	Proteína hipotética	49/67	<i>S. leeuwenhoekii</i> WP_029381414
12	143	Proteína hipotética	83/90	<i>S. lydicus</i> WP_046928240
13	180	Proteína de membrana	97/97	<i>S. azureus</i> GAP50205
14	301	Piridoxina biossíntese glutamina amidotransferase, synthase subunidade	97/99	<i>S. caelestis</i> KOT26627
15	202	Piridoxina biossíntese glutamina amidotransferase, glutaminase subunidade	97/98	<i>S. azureus</i> GAP50203
16	250	Regulador transcricional	77/89	<i>Thermobifida fusca</i> WP_011292519

Tabela B13 – Funções previstas do *cluster* sideroforo 1

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
1	279	Metiltransferase	85/92	<i>S. aureofaciens</i> WP_052842603
2	442	Transportador ABC Transporte de Sideróforo	80/87	<i>S. olivaceus</i> WP_031032805
3	136	Proteína hipotética	81/87	<i>S. chartreusis</i> WP_010039010
4	301	Treonina quinase	80/87	<i>S. griseorubens</i> WP_051747866
5	513	Transportador ABC Sideróforo síntese	87/89	<i>S. viridochromogenes</i> WP_003993985
6	497	Proteína hipotética	88/91	<i>S. afghaniensis</i> WP_020270904
7	210	HxIR regulador transcricional	83/88	<i>S. leeuwenhoekii</i> WP_029385820
8	402	Aminotransferase AlaT	96/98	<i>S. ghanaensis</i> WP_004981437
9	209	Proteína hipotética	89/94	<i>S. azureus</i> GAP46255
10	140	Tioesterase	92/95	<i>S. iakyus</i> WP_033306615
11	543	Serina-treonina quinase	86/90	<i>S. iakyus</i> WP_033306629

Tabela B14 – Funções previstas do *cluster* sideroforo 2

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
1	385	Proteína hipotética secretada	94/96	<i>S. afghaniensis</i> WP_037670969
2	438	transcriptional regulator	22/29	<i>Amycolatopsis methanolica</i> AIJ26057
3	489	2,4-diaminobutirato 4- aminotransferase (Biossíntese de sideróforos)	76/82	<i>S. avermitilis</i> WP_010983904
4	729	Sideróforo sintase (Rhizobactina sintase RhbC)	72/75	<i>S. afghaniensis</i> WP_020274560
5	361	Sideróforo sintase (Aciltransferase RhbD)	72/77	<i>S. afghaniensis</i> WP_020274561
6	658	Helicase ATP-dependente	98/98	<i>S. viridochromogenes</i> WP_003993441

7	259	LexA repressor	99/99	<i>S. afghaniensis</i> WP_020274563
8	186	Regulador transcricional NrdR Ribonucleotídeo reductase	98/97	<i>S. chartreusis</i> WP_029180822
9	959	Ribonucleotídeo reductase (dependente de coenzima B12)	98/98	<i>S. chartreusis</i> WP_010034867

Tabela B15 – Funções preditas do *cluster* Sideróforo 3 homólogo a Desferrioxamina B

ORF	No. a.a.	Função Proposta	%Identidade/ Similaridade	Homólogo e Origem No. Acesso
1	605	Glucosamine--fructose-6-phosphate aminotransferase (Glutamina amidotranferase)	97/98	<i>S. viridochromogenes</i> WP_004001033
2	85	Proteína Hipotética	99/100	<i>S. afghaniensis</i> WP_037673904
3	71	Fosfoesterase	93/94	<i>S. chartreusis</i> WP_010047880
4	293	Regulador da via Tat	89/93	<i>S. iakyrus</i> WP_033308053
5	534	Beta-N-acetilhexosaminidase	89/93	<i>S. chartreusis</i> WP_050810337
6	590	lucA/lucC Sideróforo sintase Desferrioxamine E	85/91	<i>S. yerevanensis</i> WP_051744879
7	176	Acetiltransferase Sideróforo sintase Desferrioxamine E	83/88	<i>S. viridochromogenes</i> WP_004001040
8	411	Lisina – Ornitina Monooxygenase Sideróforo Síntese	94/96	<i>S. afghaniensis</i> WP_037669866
9	480	L-2,4-Diaminobutirato descarboxilase Sideróforo Síntese	95/97	<i>S. afghaniensis</i> EPJ38541
10	284	Ferredoxine reductase Proteína de interação com Sideroforos (ViuB)	94/96	<i>S. azureus</i> GAP50066
11	365	ABC transportador Desferrioxamine E	93/96	<i>S. afghaniensis</i> WP_020273314
12	665	Glicosiltransferase	71/81	<i>S. zinciresistens</i> WP_007494653
13	386	Acil CoA Desidrogenase	93/96	<i>S. svicens</i> WP_007382220
14	316	Hidroximetilglutaril-CoA liase	94/95	<i>S. afghaniensis</i> WP_037669874
15	638	Carbamoil fosfato sintase Biotina carboxilase	95/97	<i>S. viridochromogenes</i> EFL32313
16	538	Metilcrotonil-CoA carboxilase carboxil transferase	98/98	<i>S. afghaniensis</i> WP_020273323
17	202	Regulador transcricional TetR	74/78	<i>S. niveus</i> WP_023541556
18	385	Acil Co A desidrogenase	97/99	<i>S. coelicolor</i> WP_011028577
19	295	Acil-CoA tioesterase II	60/69	<i>Amycolatopsis thermoflava</i> WP_027934992
20	265	Fosfatase	97/98	<i>S. chartreusis</i> WP_010046952
21	66	Regulador transcricional PucR	40/53	<i>S. acidiscabies</i> WP_050987331

Tabela B16 – Funções preditas do *cluster* Híbrido Melanina - Nucleosídeo

ORF	No. a.a.	Função Proposta	%Ident./ Simil.	Homólogo e Origem No. Acesso
1	287	tyrosinase	86/92	<i>S. azureus</i> GAP52171

2	176	tyrosinase co-factor protein	76/79	<i>S. azureus</i> GAP52170
3	227	hypothetical protein	32/41	<i>S. griseoaurantiacus</i> WP_040895432
4	71	Tat pathway signal protein	32/48	<i>Arthrobacter</i> sp. M2012083 WP_017200212
5	208	transposase	51/59	<i>S. anulatus</i> WP_030595985
6	513	transposase, partial	77/82	<i>S. anulatus</i> WP_051860463
7	138	hypothetical protein	61/71	<i>S. fradiae</i> WP_050363655
8	87	hypothetical protein	33/42	<i>S. sulphureus</i> WP_019544511
9	524	hypothetical protein	57/68	<i>S. atratus</i> WP_051956610
10	381	hypothetical protein	56/64	<i>Microbispora rosea</i> WP_030507203
11	393	Undecaprenyl-phosphate alpha-N-acetylglucosaminyl 1-phosphate transferase	27/42	<i>Collinsella aerofaciens</i> CUP27126
12	362	nucleoside-diphosphate sugar epimerase	64/76	<i>Streptosporangium roseum</i> WP_037935839
13	426	UDP-N-acetyl-D-glucosamine dehydrogenase	84/90	<i>S. nodosus</i> WP_052454169
14	426	polysaccharide biosynthesis protein	14/27	<i>Pseudonocardia dioxanivorans</i> WP_013674047
15	309	GDP-mannose 4,6-dehydratase	84/90	<i>S. nodosus</i> WP_043442304
16	385	glycosyl transferase	80/86	<i>S. nodosus</i> WP_052454170
17	455	hypothetical protein	41/50	<i>S. scabiei</i> WP_013001840
18	169	alanine acetyltransferase	31/42	<i>Mycobacterium smegmatis</i> WP_036452820
19	222	GlcNAc-PI de-N-acetylase	24/37	<i>Streptosporangium roseum</i> WP_031160900
20	232	hypothetical protein	30/48	<i>Mycobacterium rhodesiae</i> WP_041302775
21	386	diguanylate cyclase	21/33	<i>Corynebacterium jeikeium</i> WP_034972532
22	155	acetyltransferase	44/63	<i>S. bingchenggensis</i> WP_014181152
23	273	hypothetical protein	20/35	<i>Mycobacterium gastri</i> WP_036419510
24	254	SAM-dependent methyltransferase	25/36	<i>Nocardiopsis kunsanensis</i> WP_026116136
25	399	glycosyl transferase family 1	15/29	<i>Bifidobacterium pseudocatenulatum</i> WP_004221867
26	152	Menaquinone biosynthesis methyltransferase	26/34	<i>Saccharothrix espanaensis</i> CCH35254
27	333	UDP-glucose epimerase	43/55	<i>Rhodococcus fascians</i> WP_037173067
28	293	hypothetical protein	29/44	<i>Mycobacterium gastri</i> WP_036419510
29	315	hypothetical protein	28/42	<i>Nitriliruptor alkaliphilus</i> WP_052669737
30	289	glutaminy-peptide cyclotransferase	71/80	<i>S. atratus</i> WP_051956599