Study on the quorum sensing mechanism in a marine bacterium *Pseudoalteromonas* sp. 520P1

DANG TRAN HOANG

A dissertation submitted to

Kochi University of Technology

in partial fulfillment of the requirements for the degree of

Doctor of Engineering

Graduate School of Engineering

Kochi University of Technology

Kochi, Japan

September, 2015

Study on the quorum sensing mechanism in a marine bacterium *Pseudoalteromonas* sp. 520P1 DANG TRAN HOANG

TABLE OF CONTENTS

ABTRACT 6

ACKNOWLEDGEMENTS 11

CHAPTER I General Introduction 12

1.1. Basics of quorum sensing systems 12

1.2. Quorum sensing in Gram-negative bacteria 13

1.2.1. LuxI/LuxR quorum sensing systems 13

1.2.2. AHL synthase 14

1.2.3 Transcriptional receptor protein LuxR 17

1.2.4 Extra LuxR-type proteins 19

1.2.5 AHL-degrading enzyme 20

1.3. Violacein production and AHL-dependent quorum sensing 22

CHAPTER II Cloning the *luxI* gene from *Pseudoalteromonas* sp. 520P1 by degenerate PCR 24

2.1. Introduction 24

2.2. Materials and Methods 24

2.2.1. GenBank/DDBJ accession numbers 24

2.2.2. Baterial strains and reagents 25

2.2.3. Homology comparison of LuxI family proteins 25

2.2.4. Design of degenerate primers 25

2.3. Results 26

2.3.1. Divergence of LuxI family protein 26

2.3.2. Conserved domain in LuxI family protein 27

2.3.3. Degenerate PCR 29

2.4. Discussion 31

CHAPTER III Genome Sequencing of *Pseudoalteromonas* sp. 520P1 No. 412 32

3.1. Introduction 32

3.2. Materials and Methods 32

3.2.1. GenBank/DDBJ accession numbers 32

3.2.2. Sequencing of the whole genome of strain 520P1 No. 412 33

3.2.3. Phylogenetic analysis 33

3.3. Results 33

3.3.1. *luxI* and *luxR* gene homologues in the genome of strain 520P1 33

3.3.2. Homology comparison of LuxI family proteins in strain 520P1 and reference strains 35

3.3.3. Homology comparison of LuxR family proteins in strain 520P1 and reference strains 37

3.4. Discussion 41

CHAPTER IV Cloning and expression of the *PalI* gene from *Pseudoalteromonas* sp. 520P1 in *E. coli* 44

4.1. Introduction 44

4.2. Materials and Methods 44

4.2.1. Bacterial strains, growth conditions and plasmids 44

4.2.2. Cloning of *palII* gene of strain 520P1 No. 412 45

4.2.3. Construction of the expression plasmid for *E. coli* 45

4.2.4. Expression of *palII* gene homologue 46

4.2.5. Extraction of AHLs 46

4.2.6. Plate assay of AHLs 46
4.2.7. TLC assay of AHLs 47
4.3. Results 47
4.3.1. Cloning of *palI* gene from strain 520P1 47
4.3.2. Expression of cloned *palI* gene in *E. coli* and detection of AHL activity 48
4.3.3. Analysis of AHLs by TLC 50
4.4. Discussion 51
CHAPTER V Genome sequencing of a variant of *Pseudoalteromonas* sp. 520P1 52
5.1. Introduction 52
5.2. Materials and Methods 53
5.3. Results
5.3.1 Whole genome analysis of strain 520P1 No. 423 53
5.3.2 Comparison with the genome of strain 520P1 No. 412 53
5.4. Discussion 53
CHAPTER VI. Conclusions 55
REFERENCES 56
ABBREVIATIONS 63
APPENDIXES 64
REAGENTS PREPARATION 88

ABSTRACT

Quorum sensing is a mechanism of intercellular bacterial communication in which specific signal molecules called autoinducers regulate the expression of genes for the production of secondary metabolites. Bacterial cells constitutively produce and respond to the autoinducers, depending on the bacterial density. In Gram-negative bacteria, autoinducers of quorum sensing systems are *N*-acylhomoserine lactones (AHLs). The AHL-based quorum sensing system was found for the first time in a marine bioluminescence bacterium *Vibrio fisheri*. This type of quorum sensing contains two main components. AHL synthase (usually known as LuxI or LuxI homologue) is responsible for AHL synthesis. The second component is a regulatory protein (usually known as LuxR or LuxR homologue) which binds to AHL and promotes transcription of the target gene.

Violacein, a purple pigment produced by several species of terrestrial and marine bacteria, is a secondary metabolite that possesses various biological properties, such as antibacterial, antiviral, antitrypanosomal and antitumor activities. These bioactive properties offer the possibility to use violacein for medical applications.

It has been reported that the violacein production of *Pseudoalteromonas* sp. 520P1 is regulated by *N*-acylhomoserine lactone-mediated quorum sensing. The gene cluster responsible for violacein biosynthesis in *Pseudoalteromonas* sp. 520P1 has also been characterized. However, the *luxI* gene and *luxR* gene in strain 520P1 have not been identified so far. Identification of these genes is crucial to extend our knowledge of the whole regulatory mechanism of the violacein biosynthesis in *Pseudoalteromonas* sp. 520P1. In this study, we investigated the protein components (LuxI and LuxR

homologues) of AHL-dependent quorum sensing system in *Pseudoalteromonas* sp. 520P1 and expressed the cloned *luxI* in *Escherichia coli*.

PART-I

Cloning the *luxI* gene from *Pseudoalteromonas* sp. 520P1 by degenerate PCR

Degenerate PCR method has been widely used to amplify DNA when only protein sequences of the genes of interest are identified, or the isolation of similar genes from related organisms is intended.

It has been reported that the LuxI family proteins (AHL synthases) have been found in more than 150 species of Proteobacteria. There is a great diversity among the AHL synthases in their production of AHLs. The sequence conservation between any two members of the AHL synthase family is typically greater than 40% similarity.

In this study, we used degenerate primers to amplify the partial of *luxI* gene of strain 520P1 based on the nucleotide sequences deduced from the conserved protein sequences between LuxIs from different bacterial strains.

The result of degenerate PCR indicated that the partial regions of *luxI* gene homologue in *Pseudoalteromonas* sp. 520P1 could not be amplified using degenerated primers. The reason for this failure might come from the low homology of LuxIs at nucleotide level. Therefore, to identify and characterize the *luxI* gene homologue in strain 520P1, it is essential to sequence the whole genome of *Pseudoalteromonas* sp. 520P1.

PART-II

Genome Sequencing of *Pseudoalteromonas* sp. 520P1 No. 412

The genome of strain 520P1 No. 412 was sequenced on an Illumina Hiseq 2000 system and one *luxI* and five *luxR* homologues were identified. Based on the amino acid

sequences, we characterized one LuxI and five LuxR protein homologues in *Pseudoalteromonas* sp. 520P1 named PalI and PalR1 to PalR5, respectively. The properties of PalI and five PalRs provided the basis for understanding the components of *N*-acylhomoserine lactone-dependent quorum sensing system in this marine bacterium.

We compared the amino acid sequence of PalI with those of LuxI homologues from different bacterial species. Based on the structure of EsaI, an AHL synthase well-defined using X-ray crystallography, we identified six amino acid regions that share significantly conserved homology in all aligned sequences. The results suggest that PalI is an AHL synthase and has a close relationship with LuxI from *Pseudoalteromonas atlantica* T6c among LuxI homologues in *Pseudoalteromonas*.

We identified five LuxR homologues (PalR1 to PalR5) in the genome of the strain 520P1, in which one pair of PalI and PalR3 were located adjacently on the scaffold 19. When compared with LuxRs from other *Pseudoalteromonas* species, all sequences shared low homology in the AHL-binding domain (N-terminus) but retained some conserved amino acid sequences in the DNA-binding domain (C-terminus). The divergence of N-terminal regions in LuxR homologues of strain 520P1 was consistent with the general characteristics of LuxR family proteins from different species, including *Pseudoalteromonas* species.

Among LuxR homologues of strain 520P1, we found PalR5 shared 74 % and 77 % homology with LuxRs from *Pseudoalteromonas citrea* DSM8771 and *Pseudoalteromonas luteoviolacea*, respectively. On the other hand, phylogenetic analysis showed that other LuxR homologues of strain 520P1, PalR1, PalR2, PalR3 and PalR4, were not in the same branch with PalR5 and phylogenetically separated from each other. The reason for this could be explained when we assume that palR genes have been

acquired from different sources during the evolutionary process of *Pseudoalteromonas* sp. 520P1 probably by horizontal gene transfer. Since the cognate LuxI homologues of these PalRs except PalR3 have not been identified, it is most likely that they are LuxR solos. In the strain 520P1, these LuxR solos might respond to endogenous and exogenous signals produced by neighboring bacteria and then control a set of genes.

PART-III

Cloning and expression of the *PalI* gene from *Pseudoalteromonas* sp. 520P1 in *E. coli*

To examine whether the *palI* gene product indeed has AHL synthase activity, we tried to clone this gene and express in *E. coli* BL21 (DE3). We successfully cloned the *palI* gene and expressed in *E. coli* BL21 (DE3). Using AHL bioassay coupled to TLC plate, two types of AHL were detected in the extract of the expression culture of recombinant *E. coli* harboring *palI* gene. Two spots of AHL on the TLC plate were similar in mobility to those of strain 520P1 No. 412.

Bioassay of AHL on a TLC plate in this study demonstrated that PalI can produce in *E. coli* two types of AHL similar to those in strain 520P1. Therefore, possible involvement of PalI in the regulation of violacein synthesis in strain 520P1 is suggested if we can assume that PalI is the only AHL synthase in strain 520P1.

PART-IV

Genome sequencing of a variant of *Pseudoalteromonas* sp. 520P1

Pseudoalteromonas sp. 520P1 No. 423, a variant strain that was isolated from the culture of strain 520P1, has an ability of highly stable production under agitated culture conditions. In the previous study, it was also hypothesized that a mutation in the upstream

promotor region of violacein gene cluster in strain 520P1 No. 423 led to the production of violacein under agitated culture conditions. However, nucleotide sequencing showed that the upstream sequence of violacein gene cluster of strain 520P1 No. 423 was identical to that of the original strain 520P1 No. 412. Therefore, no mutation occurred in the upstream region of violacein gene cluster of strain 520P1 No. 423. To clarify the difference of *palI* and *palR* genes between strain 520P1 No. 412 and No. 423, we performed the genomic sequencing of strain 520P1 No. 423.

The results of partial genomic analysis between strain 520P1 No. 412 and No. 423 showed that no mutation existed in the nucleotide sequences of the *palI* and its promoter, five *palRs* and their promoters, violacein gene cluster and its promoter. Therefore, unique properties of strain 520P1 No. 423 in violacein production should be ascribed to the gene mutation that is not relevant to *palI*, *palR*s and violacein gene cluster.

To reveal the difference between strain 520P1 No. 412 and No. 423, we need to consider other possible reasons. One of them is the mutation in the sequences of AHL-degrading enzymes in strain 520P1 No. 423. There are two group of AHL-inactivating enzymes; AHL lactonases which hydrolyze the lactone ring in AHLs and AHL acylases which hydrolyze *N*-acyl bond and release a free homoserine lactone and a fatty acid. We suggest that the identification of these enzymes might clarify the ability of producing violacein under agitated culture conditions of strain 520P1 No. 423.

Keywords: *Pseudoalteromonas*; violacein; *N*-acylhomoserine lactone; quorum sensing

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincerest gratitude to my supervisor, Prof. Keiichi Enomoto, for his guidance, caring and help me improve my knowledge in molecular biology and biochemistry. I appreciate all his contributions of time, ideas, and patience to make my research productive and stimulating. He has spent much time guiding me to approach research problems correctly and efficiently. I also wish to thank him for giving me the opportunity to study in Japan and experience Japanese life and culture.

For this dissertation I would like to thank my reading committee members: Associate Prof. Osamu Ariga, Associate Prof. Sakae Hosikawa, Prof. Kenzo Sato, Prof. Masaki Sazuka and Associate Prof. Yasunori Matsumoto for their constructive comments and suggestions in my research.

I would like to thank all members of the Prof. Enomoto's Lab for their help, support and friendship. Special thanks to Mr. Ryota Fujikawa, Ms. Tomomi Omosako, Mr. Kentarou Yotsumoto, Mr. Shinya Komatsu, Mr. Hideyuki Masuda for their warm collaborations during my research.

I gratefully acknowledge the funding source that made my PhD work possible. I was funded by the Special Scholarship Program of Kochi University of Technology (KUT) for 3-years Doctoral Course. I am grateful to staffs of International Relations Division (IRC) of KUT who kept us organized and was always ready to help.

Finally, I would like to thank my family for all their love and encouragement.

Dang Tran Hoang
Kochi University of Technology
September 2015

CHAPTER I

General Introduction

1.1 Basics of the Quorum Sensing System

Quorum sensing is a mechanism of intercellular bacterial communication in which specific signal molecules called autoinducers regulate the expression of genes for the production of secondary metabolites. Bacterial cells constitutively produce and respond to the autoinducers, depending on the bacterial density. This communication mechanism has been found in various bacterial strains. Autoinducers are produced in parallel with the growth of bacteria and then excreted by diffusion to the outside environment.

When the concentration of autoinducers reaches a critical level (usually mentioned as threshold level), they bind to receptor proteins to become the complexes of autoinducer and receptor protein. These complex then bind to specific DNA region that located very near the region of corresponding target genes. Consequently, this interaction results in the transcription of target genes and helps bacteria to respond to the changes of their living environment (Miller and Bassler 2001; Ng and Bassler 2009; Galloway et al. 2011).

It has been proven that bacteria use quorum sensing to control certain types of phenotypic effects such as biofilm formation, virulence, and antibiotic resistance etc. Quorum sensing is not just a control mechanism involving many functional processes in the cells of bacteria, it also occurs between bacteria of the same species and between different species of bacteria.

Autoinducers of quorum sensing systems derive from different sources. In general, signaling molecules are oligopeptides in Gram-positive bacteria, *N*-acylhomoserine

lactones (AHLs) in Gram-negative bacteria, and autoinducer-2 (AI-2), a new group of autoinducers recently identified and characterized, in both Gram-negative and Gram-positive bacteria (Miller and Bassler 2001).

1.2 Quorum Sensing in Gram-negative bacteria

1.2.1 LuxI/LuxR quorum sensing systems

In Gram-negative bacteria, signal molecules of quorum sensing systems are *N*-acylhomoserine lactones (AHLs) (McClean et al. 1997; Morohoshi et al. 2010).

The AHL-based quorum sensing system was found for the first time in a marine bioluminescence bacterium *Vibrio fisheri* (Miller and Bassler 2001). This type of quorum sensing contains two main components. The first is an AHL synthase (usually known as LuxI or LuxI homologue) is responsible for AHL synthesis. LuxI catalyzes the enzymatic reaction to form AHLs using acyl-CoA derivatives as the main acyl chain donor and S-adenosylmethionine (SAM) as the source for the homoserine lactone moiety via an acyl carrier protein (Hanzelka et al. 1996; Williams 2007; Churchill and Chen 2011; Decho et al. 2011).

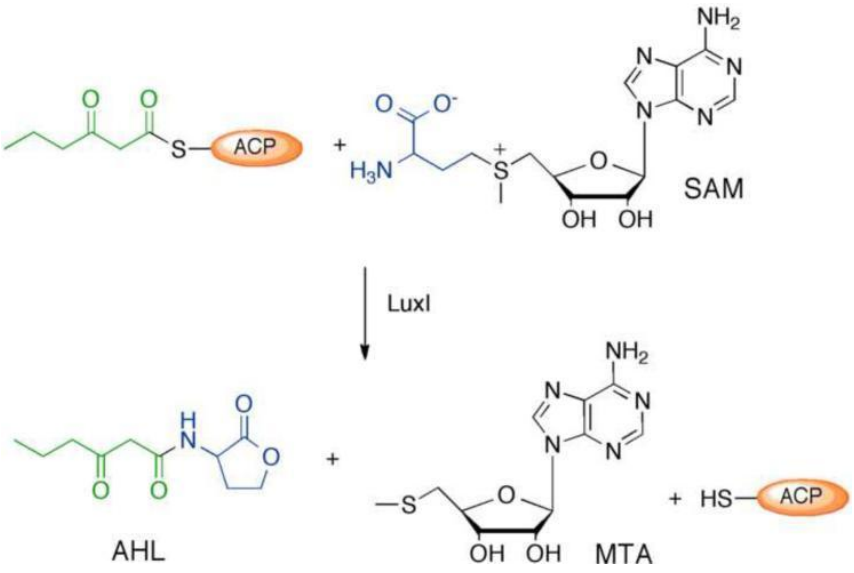


Fig.1 Enzymatic reaction catalyzed by LuxI (Fast and Tipton 2012)

The second component of AHL-based quorum sensing systems is a regulatory protein (usually known as LuxR or LuxR homologue) which promotes transcription of the target gene.

LuxI and LuxR homologues have been found in various Gram-negative bacteria with a diversity of different AHLs. In general, *luxI* and *luxR* homologue genes exist in cognate pairs in which LuxR only bind to AHL molecule synthesized by cognate LuxI (Lerat and Moran 2004; Galloway et al. 2011).

At present, more than 70 different Gram-negative bacteria have been reported to possess the LuxI/LuxR-type quorum sensing system (Czajkowski and Jafra 2009).

1.2.2 AHL synthase

Biosynthesis of AHLs has been found in more than 90 bacterial strains. AHLs are synthesized by enzymes known as AHL synthase. These enzymes are divided into three separate groups, including LuxI, HdtS, and LuxM-type protein families.

The LuxI-type proteins are common group of AHL synthase and have been widely studied in many different microorganisms. Until now, LuxI-type proteins have been identified in more than 150 species of Proteobacteria.

In general, each AHL molecule is composed of a homoserine lactone ring (HSL) with an acyl chain. The structures of AHLs were shown in **Fig. 2**. AHLs produced by different bacteria differ in the length of the R group (the difference in the length of acyl side chains and the oxidation or hydroxylation at C3 position). The number of carbon atoms in the acyl side chains generally varies between 4 and 16 carbons.

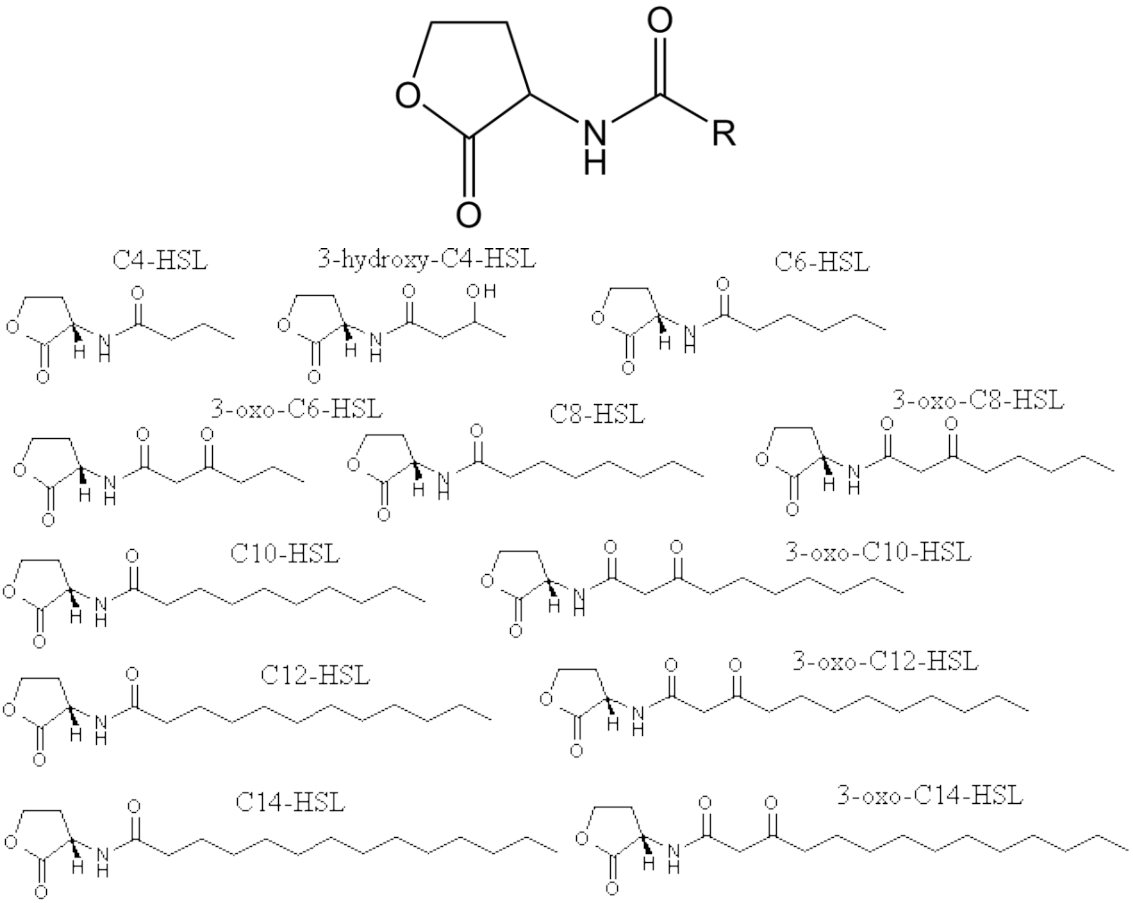


Fig.2 Structures of AHLs

(https://en.wikipedia.org/wiki/N-Acyl_homoserine_lactone;
http://www.nottingham.ac.uk/quorum/AHLs.htm)

Some bacterial strains such as *V. fischeri* and *Agrobacterium tumefaciens* synthesize a single type of AHL, 3-oxo-C6-HSL for *V. fischeri* or 3-oxo-C8-HSL for *A. tumefaciens*, respectively (**Table.1**). These microorganisms usually contain a single AHL synthase gene that is responsible for the production of a certain type of AHL. In addition, analysis of bacterial genomes has revealed that some bacteria may have more than one AHL synthase gene. Therefore, these bacteria are able to produce a mixture of AHLs. *Pseudomonas aeruginosa*, for example, has two LuxI-type AHL synthases, LasI and RhlI. These enzymes synthesize 3-oxo-C12-HSL and C4-HSL, respectively.

Table 1. LuxI-type AHL synthases and AHL signals*		
LuxI	*Vibrio fischeri*	3-oxo-C6-HSL
EsaI	*Pantoea stewartii*	3-oxo-C6-HSL
TraI	*Agrobacterium tumefaciens*	3-oxo-C8-HSL
LasI	*Pseudomonas aeruginosa*	3-oxo-C12-HSL
PhzI	*Pseudomonas fluorescens*	3-hydroxy-C6/C8-HSL
RhlI	*Pseudomonas aeruginosa*	C4-HSL
CviI	*Chromobacterium violaceum*	C6-HSL
BmaI1	*Burkholderia mallei*	C8/C10-HSL
CerI	*Rhodobacter sphaeroides*	7-cis-C14-HSL

*Churchill and Chen 2011

The secondary structure of AHL synthase family has also been determined. These enzymes have a single domain of approximately 205 amino acid residues in length. The degree of sequence conservation between any two members of the AHL synthase family is usually larger than 40% similarity. Despite the low degree of similarity, there remains a typical sequence motif of entirely conserved amino acid residues that validates the LuxI-type AHL synthase family.

In general, the N-terminus of LuxI-type AHL synthases contains the most conserved amino acids (amino acid residues 1-100). In this region, eight invariant residues critical for AHL synthesis were identified. They are Arg24, Phe28, Trp34, Asp45, Asp48, Arg68, Glu97, and Arg100 (the number of amino acid residues is based on the structure of EsaI (PDB ID: 1KZF), a well-studied AHL synthase from *Pantoea stewartii* subsp. *stewartii* and LasI (PDB ID: 1RO5) from *P. aeruginosa*) (**Fig.3**). The high degree of sequence conservation in the N-terminus suggested an important role in catalytic

process and in the binding of SAM. In contrast, the C-terminus of AHL synthases is less conserved and appears to be less important for activity in some types of AHL synthase such as RhlI in *P. aeruginosa*. The role of C-terminus is not really clear and need to be studied further. Nevertheless, this region is involved in recognition of the most variable region of the acyl-ACP substrate, the acyl chain, which elucidates the higher degree of variability in this region. As a consequence, LuxI-type AHL synthase family share similar structures and mechanisms of AHL synthesis (Churchill and Chen 2011).

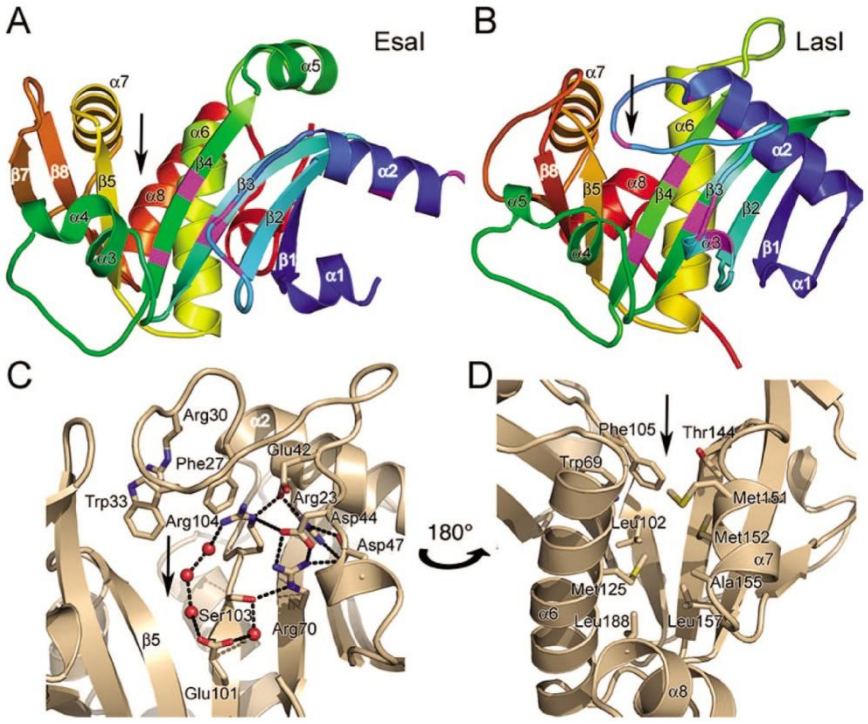


Fig.3 Structure of the AHL synthases EsaI and LasI (Churchill and Chen 2011)

A, EsaI from *Pantoea stewartii* subsp. *stewartii*; B, structure of LasI from *Pseudomonas aeruginosa*; C, close-up view of the LasI active site residues; D, close-up view of the LasI acyl-chain binding pocket.

1.2.3 Transcriptional receptor protein LuxR

In AHL-based quorum sensing systems, AHLs are recognized by two types of receptors. The first type is LuxR and its cognate AHL, components of the quorum sensing mediated bioluminescence in *V. fischeri*. The LuxR-type proteins promote transcription

of target genes by forming direct contact with DNA sequences in the vicinity of the promoter region of target genes. This DNA binding activity is regulated by AHLs.

In the second way, membrane-bound sensor kinases such as LuxN and LuxQ of *Vibrio harveyi* are responsible for the binding with AHLs and promote the transcription of target genes by means of a phosphorelay mechanism.

Until now, LuxR-type proteins are found more common and form a family of LuxR protein. LuxR family proteins are consisted of approximately 250 amino acid and have two main functional domains. The N-terminal domain (NTD) is responsible for binding with AHL. The C-terminal domain (CTD) is responsible for the DNA binding.

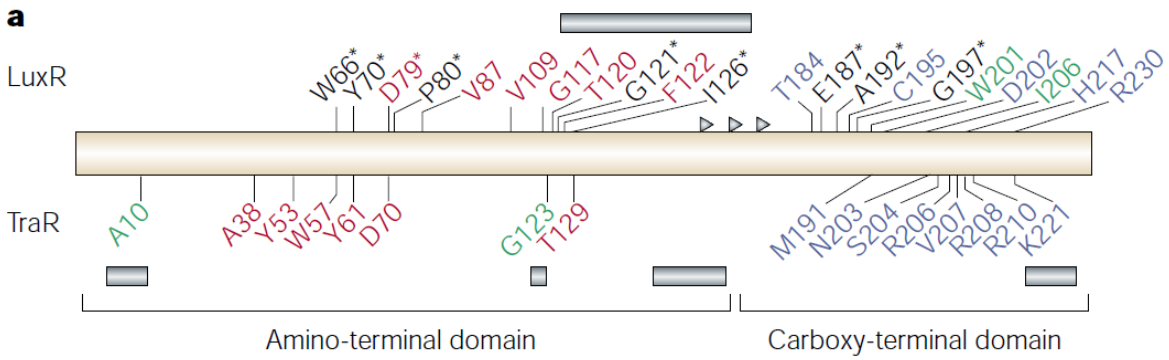


Fig.4 Functional domains of LuxR family proteins (Fuqua and Greenberg 2002)

Residues labeled with an asterisk are conserved in most members of the LuxR protein family; residues involved in interactions with the relevant AHL are in red; residues involved in DNA binding are in blue; residues involved in transcription activation are in green.

The LuxR-type proteins can be transcription activators in which when binding to the target DNA region, the complex of LuxR-type protein and AHL promotes the transcriptional activation of the target genes. The properties of the transcriptional receptor LuxR-type proteins have been studies in LuxR (*V. fischeri*, in bioluminescence), LasR/RhlR (*P. aeruginosa*, virulence factor expression and biofilm formation), and TraR (*A. tumefaciens*, oncogenic Ti plasmid replication and conjugal transfer).

However, a few LuxR-type proteins are known to have transcriptional repression activity. Among them, EsaR (*Pantoeastewartii*, in exopolysaccharide production) binds to the target DNA region of the promoter region in the absence of AHLs, which results in the repression of target genes. The precise mechanism of EsaR has not yet been elucidated. However, it is believed that the presence of AHL has changed the DNA binding domain of EsaR thereby inhibiting the expression of the target gene (Whitehead et al. 2001; Bottomley et al. 2007; Churchill and Chen 2011).

1.2.4 Extra LuxR-type proteins

A typical AHL-dependent quorum sensing system commonly contains two components, LuxI and LuxR protein families.

In general, the location of cognate *luxI* and *luxR* genes are usually adjacent to each other in the genome. However, the analysis of complete genome sequences in different microorganisms has revealed that various strains possess LuxRs that do not have a coupled LuxI protein. These LuxRs have been considered as individual LuxRs or LuxR solos. LuxR solos are present widely both in species of Proteobacteria group that possesses a complete AHL-mediated quorum sensing system and in species that do not. In addition, it has been reported that several bacterial genomes contain more than one LuxR solo (Case et al. 2008; Patankar and González 2009; Subramoni and Venturi 2009).

LuxR solos generally possess the typical protein structure of LuxR-type protein family. However, there are considerable variations in the length of amino acid sequence and some proteins lack conserved amino acids within the AHL-binding domain (Fuqua and Greenberg 2002).

It has been reported that *Sinorhizobium meliloti*, *Rhizobium leguminosarum bv. viciae* and *Pseudomonas aeruginosa* contain LuxR solos, ExpR, BisR and QscR,

respectively. The role of these LuxR solos in relation with AHL-binding ability and quorum sensing mechanism of bacteria still remains to be clarified (Subramoni and Venturi 2009; Oinuma and Greenberg 2011).

Table 2. Orphan LuxR homologues in Gram-negative bacteria*

Organism	Orphan LuxR homologue	Cognate LuxI/LuxR pair
Pseudomonas aeruginosa	QscR	LasI/LasR
Sinorhizobium meliloti	ExpR	SinI/SinR
Rhizobium leguminosarum bv. viciae	BisR	CinI/CinR
Agrobacterium tumefaciens	TrlR	TraI/TraR
Salmonella enterica	SdiA	none
E. coli	SdiA	none
Brucella melitensis	VjbR, BlxR	none
Burkholderia pseudomallei	BpmR4, BpmR5	BmlI/BmlR BpmI2/BpmR

*Patankar and González 2009

1.2.5 AHL-degrading enzyme

In many pathogenic bacteria, the quorum sensing controls virulence factor synthesis. It has been proven that the inactivation of the quorum sensing is controlled by AHL-degrading enzymes.

Interference in the quorum sensing mechanism can be achieved in different ways. First, many natural substances can disturb the signal perception by imitating AHLs structure. The AHL analogues block the AHL receptor (regulator) protein and then prevent activation of the target gene expression (Czajkowski and Jafra 2009).

Until now, AHL-degrading enzymes have been found in different bacteria. The structure of AHLs suggests four different ways for degrading AHLs. As shown in **Fig.5**, positions 1 and 2 are for the degradation of the homoserine lactone ring catalyzed by lactonase or decarboxylase. Positions 3 and 4 are the locations where acylase or

deaminase degrades AHL into a homoserine lactone and a free fatty acid moiety (Dong and Zhang 2005).

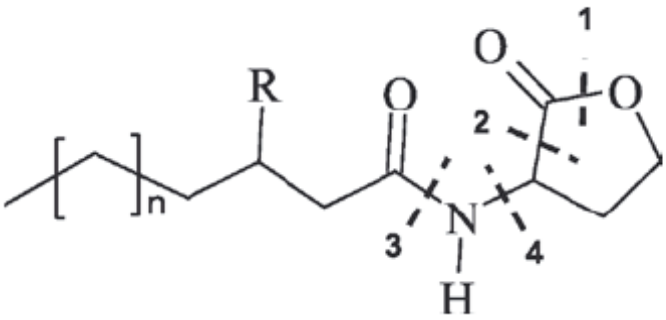


Fig.5 The possible ways of enzymatic degradation of AHLs (Czajkowski and Jafra 2009)

1, lactonase; 2, decarboxylase; 3, acylase; 4, deaminase

There are two main groups of AHL-degrading enzymes well characterized, including AHL lactonases (EC 3.1.1.81) and AHL acylases (EC 3.5.1.97).

AHL lactonases hydrolyze the lactone ring in the homoserine moiety of AHLs, while AHL acylases hydrolyze the amide bond between the acyl side chain and the homoserine lactone to release the free fatty acid and the homoserine lactone.

Table 3. AHL-degrading enzymes in various bacterial species*

Organism	AHL-degrading protein	Enzymatic activity
Pseudomonas aeruginosa	PvdQ, QuiP	AHL acylase
Agrobacterium tumefaciens	AttM, AiiB	AHL lactonase
Bacillus sp.	AiiA	AHL lactonase
Bacillus thuringiensis	AiiA	AHL lactonase
Arthrobacter sp.	AhlD	AHL lactonase
Rhodococcus erythropolis	QsdA	AHL lactonase
Ralstonia eutropha	AiiD	AHL acylase
Anabaena sp.	AiiC	AHL acylase

*****Czajkowski and Jafra 2009

1.3 Violacein production and AHL-dependent quorum sensing

Violacein, a purple pigment produced by several species of terrestrial and marine bacteria, is a secondary metabolite that possesses various biological properties, such as antibacterial, antiviral, antitrypanosomal and antitumor activities (Hoshino 2011; Durán and Menck 2001; Durán et al. 2012). These bioactive properties offer the possibility to use violacein for medical applications.

Until now, it has been reported that violacein can be produced by various Gram-negative bacteria, including *Chromobacterium violaceum* (Durán and Menck 2001), *Janthinobacterium lividum* (Pantanella et al. 2007), *Pseudoalteromonas luteoviolacea* (Yang et al. 2007), and *Pseudoalteromonas* sp. 520P1 (Yada et al. 2008).

Studies on the biosysthesis of violacein have been conducted to reveal the components involved in the biosynthetic pathway. It has been reported that five enzymes, VioA to VioE, are responsible for the production of violacein in *Chromobacterium violaceum* (August et al. 2000; Morohoshi et al. 2010). These enzymes are encoded by five genes, vioA to vioE (Momen and Hoshino 2000) that form a single operon, vioABCDE (August et al. 2000). VioE is demonstrated to play an important role in synthesizing the molecular skeleton of violacein (Hirano et al. 2008).

In recent studies, Wang *et al.* (2008) reported that the production of violacein by *Pseudoalteromonas* sp. 520P1, a Gram-negative marine bacterium isolated from the Pacific coast of Japan, was also under the regulation of an AHL-dependent quorum sensing. The gene cluster responsible for violacein biosynthesis in *Pseudoalteromonas* sp. 520P1 has also been characterized (Zhang and Enomoto 2011). A cluster of five ORFs (vioABCDE) of the length of 7, 383bp has been cloned. However, the *luxI* gene and *luxR* gene in strain 520P1 have not been identified so far. Identification of these genes is crucial to extend our knowledge of the whole regulatory mechanism of the

violacein biosynthesis in *Pseudoalteromonas* sp. 520P1.

In this study, we investigated the protein components (LuxI and LuxR homologues) of AHL-dependent quorum sensing system in *Pseudoalteromonas* sp. 520P1 and expressed the cloned *luxI* in *Escherichia coli*.

CHAPTER II

Cloning the *luxI* gene from *Pseudoalteromonas* sp. 520P1 by degenerate PCR

2.1 Introduction

Degenerate PCR method has been widely used to amplify DNA in situations where only protein sequences of the genes of interest from closely related organisms are identified, or where the aim is to isolate similar genes from a variety of species. Generally, instead of using specific PCR primers with a given nucleotide sequence, degenerate PCR use mixed PCR primers designed by aligning the protein sequences of related strains to find the most conserved amino acid regions. Based on the conserved amino acid sequences, a set of degenerate PCR primers were designed using the International Union of Pure and Applied Chemistry (IUPAC) system for degenerate bases (Iserte et al. 2013).

It has been reported that the LuxI family proteins (AHL synthases) are the most widespread and best understood. Identifiable LuxI homologues have been found in more than 150 species of Proteobacteria. There is a great diversity among the AHL synthases in their production of AHLs. The sequence conservation between any two members of the AHL synthase family is typically greater than 40% similarity (Churchill and Chen 2011).

In this study, we used degenerate primers to amplify the partial of *luxI* gene of strain 520P1 based on the nucleotide sequences deduced from the conserved protein sequences between LuxIs from different bacterial strains.

2.2 Materials and Methods

2.2.1 GenBank/DDBJ accession numbers

The accession numbers of LuxI family protein sequences in DDBJ/EMBL/GenBank used in this study were P54656 (EsaI, *Pantoea stewartii* subsp.

stewartii), ABG38892 (LuxI, *Pseudoalteromonas atlantica* T6c), ERG18256 (LuxI, *Pseudoalteromonas citrea* DSM8771), WP_046355763 (LuxI, *Pseudoalteromonas luteoviolacea*), YP_206882 (LuxI, *Vibrio fischeri* ES114), WP_005423459 (LuxI, *Aliivibrio fischeri*), AAQ61751 (CivI, *Chromobacterium violaceum* ATCC 12472), P33907 (TraI, *Agrobacterium fabrum* str. C58), P54291 (RhlI, *Pseudomonas aeruginosa* PAO1), P33883 (LasI, *Pseudomonas aeruginosa* PAO1).

2.2.2 Baterial strains and reagents

Genomic DNA of *Pseudoalteromonas* 520P1 No. 412 (NBRC 107704) and *Vibrio* sp. 402W9 were used as DNA template for PCR amplification.

Ex-taq DNA polymerase for PCR amplification was purchased from Takara (Otsu, Japan).

2.2.3 Homology comparison of LuxI family proteins

Homologous sequences were multiply aligned using Clustal X2 (Larkin et al 2007). Conserved regions and gaps were determined through alignment of amino acid sequences of LuxI families.

2.2.4 Design of degenerate primers

Based on conserved amino acid regions between LuxI family proteins, the degeneracy of each amino acid and their codons was determined using the IUPAC system for degenerate bases (**Appendix 1**). Then, the nucleotide sequences of conserved amino acid sequences have been deduced. Degenerate primers were designed using the degeneracy code found in the reference IUPAC table.

Table 2.1 Degenerate bases used for PCR amplification

IUPAC nucleotide code	Base
W	A or T
R	A or G

Y	C or T
M	A or C
K	G or T

2.3. Results

2.3.1. Divergence of LuxI family protein

In the previous study**,** Churchill and Chen (2011) reported that the most conserved part of the LuxI family proteins is the N-terminal region (amino acid residues 1-100). This region contains invariant residues which are essential for enzymatic activity of AHL synthase. The high degree of sequence conservation in N-terminus suggested a role in catalysis and in the binding of the common substrate S-adenosylmethionine (SAM). Conversely, the C-terminus of LuxI is less conversed overall and appears to be less important for activity of AHL synthase. In addition, this region is involved in recognition of the most variable part of the acyl chain. Thus, the C-terminus of LuxI family proteins have the higher degree of variability than the N-terminus.

We collected the amino acid sequences of several LuxI family proteins including three LuxI homologues from *Pseudoalteromonas* strains (*P. atlantica*, *P. citrea* and *P. luteoviolacea*). The homology comparison of these sequences was performed. The result of amino acid sequence alignment showed that all LuxI homologue sequences share a low degree of homology, except several conserved amino acid residues in the N-terminus (**Fig. 2.1**).

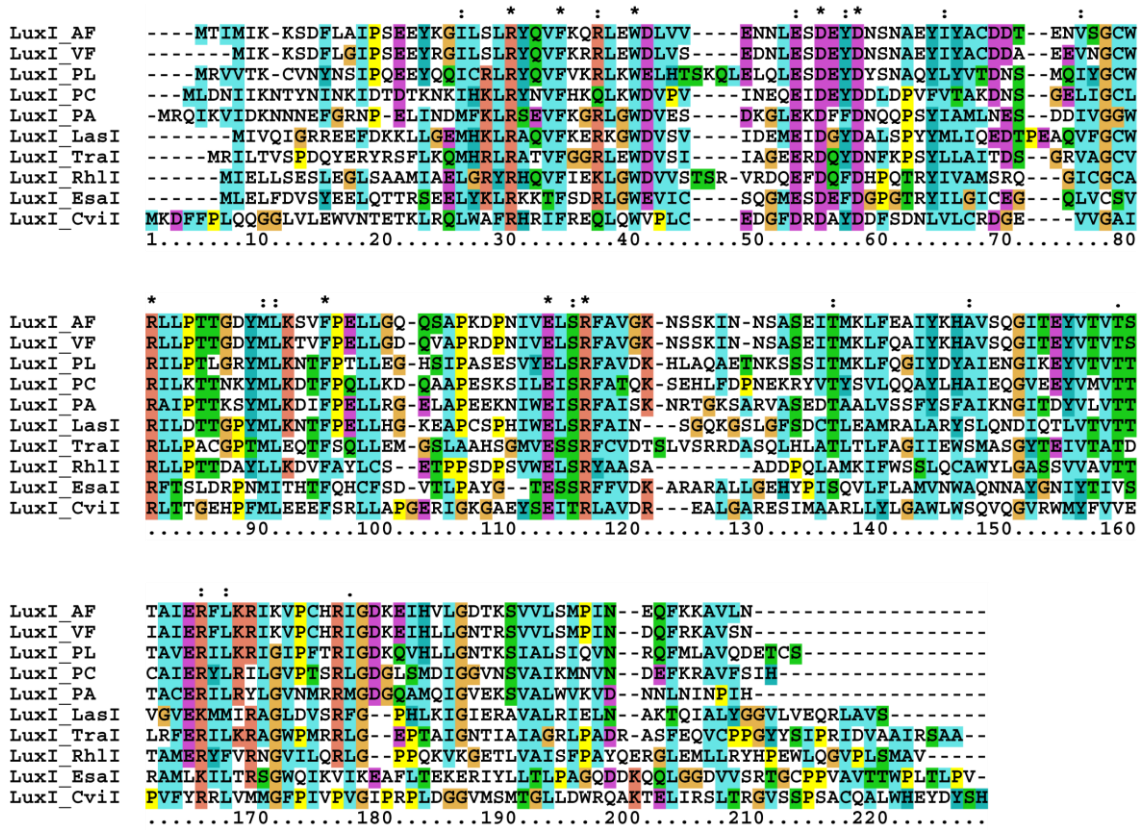


Fig.2.1 The divergence of amino acid sequences of LuxI family proteins

2.3.2. Conserved domain in LuxI family protein

The homology comparison of LuxI homologues from *P. atlantica* and *P. citrea* was performed. The result of amino acid sequence alignment showed four regions with high degree of homology. These regions were indicated as box numbers 1, 2, 3 and 4, respectively (**Fig. 2.2**).

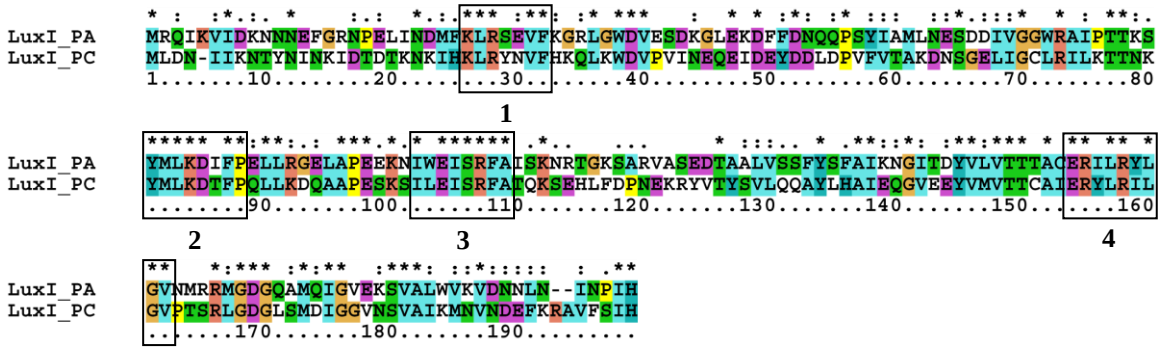


Fig.2.2 Amino acid sequence alignment of LuxI homologues.

PA: LuxI from *P. atlantica*; PC: LuxI from *P. citrea*

The homology comparison of LuxI homologues from *Pseudoalteromonas* group and CviI from *Chromobacterium violaceum* was performed. The result of amino acid sequence alignment showed low homology between LuxI homologues of *Pseudoalteromona*s strains and CviI in the four conserved regions (**Fig. 2.3**).

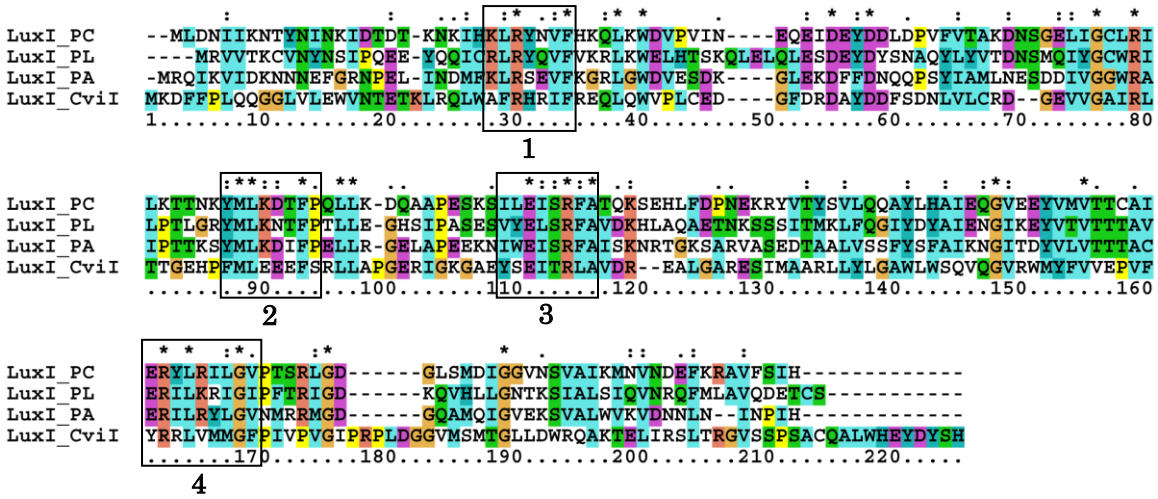


Fig.2.3 Amino acid sequence alignment of LuxI homologues.

PA: LuxI from *P. atlantica*, PC: LuxI from *P. citrea*, PL: LuxI from *P. luteoviolacea*; CviI: LuxI homologue from *C. violaceum*.

The homology comparison of LuxI homologues from *Pseudoalteromonas* strains and *Vibrio* strains was performed. The result of amino acid sequence alignment showed higher degree of homology between LuxI homologues of *Pseudoalteromonas* and *Vibrio* strains in the four conserved regions (**Fig. 2.4**).

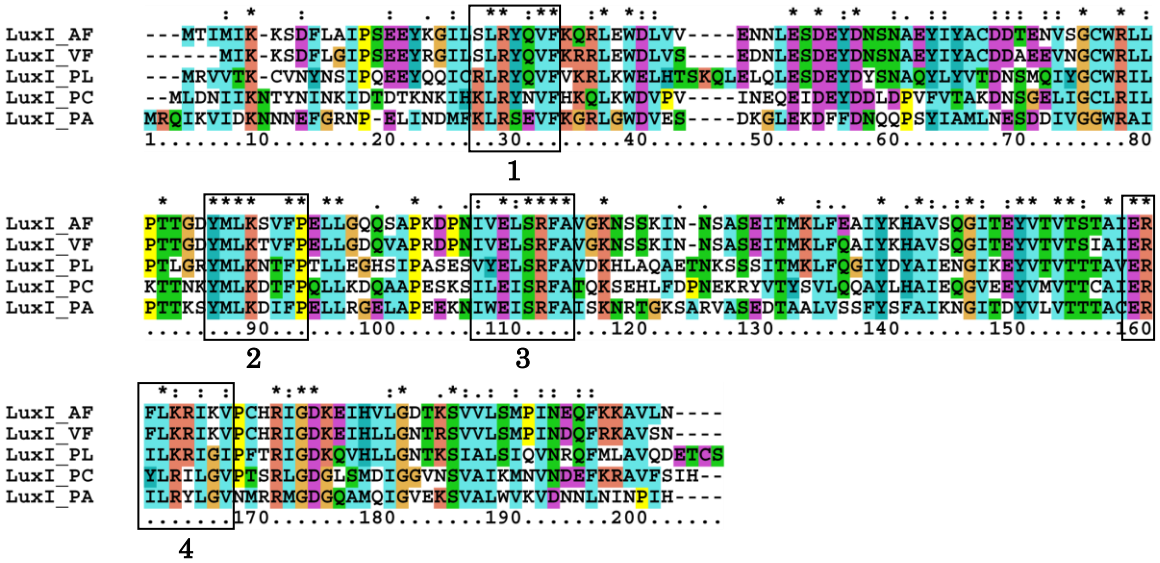


Fig.2.4 Amino acid sequence alignment of LuxI homologues.

PA: LuxI from *P. atlantica*, PC: LuxI from *P. citrea*, PL: LuxI from *P. luteoviolacea*, AF: LuxI from *A. fischeri*, VF: LuxI from *V. fischeri*

2.3.3. Degenerate PCR

Based on the amino acid sequence alignment between LuxIs from *P. atlantica* and *P. citrea*, we designed four pairs of degenerate primers to amplify four partial nucleotide sequences of *luxI* gene homologue in strain 520P1. The detailed information of degenerate primers and PCR conditions were shown in **Table 2.2** and **Table 2.3**.

Table 2.2 Synthetic degenerate oligonucleotide primers

For amplification of luxI gene homologue in strain 520P1 No. 412		
Region	**Primers**	**Sequences**
1 to 3	Fw-1-2	5’-AAACTGCGCWRYRAMGTGTTT-3’
	Rv-2-2	5’-CGCAAAGCGGCTAATTTCCMRAAT-3’
1 to 4	Fw-1-2	5’-AAACTGCGCWRYRAMGTGTTT-3’
	Rv-3-2	5’-CACGCCCAGAWWGCGCAGAWWGCGTTC-3’
2 to 4	Fw-2-2	5’-TATATGCTGAAAGATAYYTTTCCG-3’
	Rv-3-2	5’-CACGCCCAGAWWGCGCAGAWWGCGTTC-3’
3 to 4	Fw-3-2	5’-ATTYKGGAAATTAGCCGCTTTGCG-3’
	Rv-3-2	5’-CACGCCCAGAWWGCGCAGAWWGCGTTC-3’
For amplification of luxI gene homologue in strain Vibrio sp. 402W9		
1 to 3	Fw-V1-1	5’-ARMCTGCGCWRYSARGTGTTT-3’
	Rv-V3-1	5’-CGCAAAGCGGCTMAKTTCCMMAAT-3’
1 to 4	Fw-V1-1	5’-ARMCTGCGCWRYSARGTGTTT-3’
	Rv-V4-1	5’-CACKYYMAKRYRKYKCAGAAWGCGTTC-3’
2 to 4	Fw-V2-1	5’-TATATGCTGAAARMYRTKTTTCCG-3’
	Rv-V4-1	5’-CACKYYMAKRYRKYKCAGAAWGCGTTC-3’
3 to 4	Fw-V3-1	5’-ATTKKGGAAMTKAGCCGCTTTGCG-3’
	Rv-V4-1	5’-CACKYYMAKRYRKYKCAGAAWGCGTTC-3’

Table 2.3 PCR conditions for amplifying the partial sequences of *luxI* gene homologue in strain 520P1

Temperature (^{o}C)	Time	Cycles
95	5 min	1
95	15 sec	45
40	30 sec	
72	30 sec	
72	5 min	1
4	∞	1

Due to high homology in the four conserved amino acid regions between LuxIs from *Pseudoalteromonas* and *Vibrio* strains, we decided to choose genomic DNA of the *Vibrio* sp. 402W9, a bioluminescent bacterium, as a positive control for degenerate PCR.

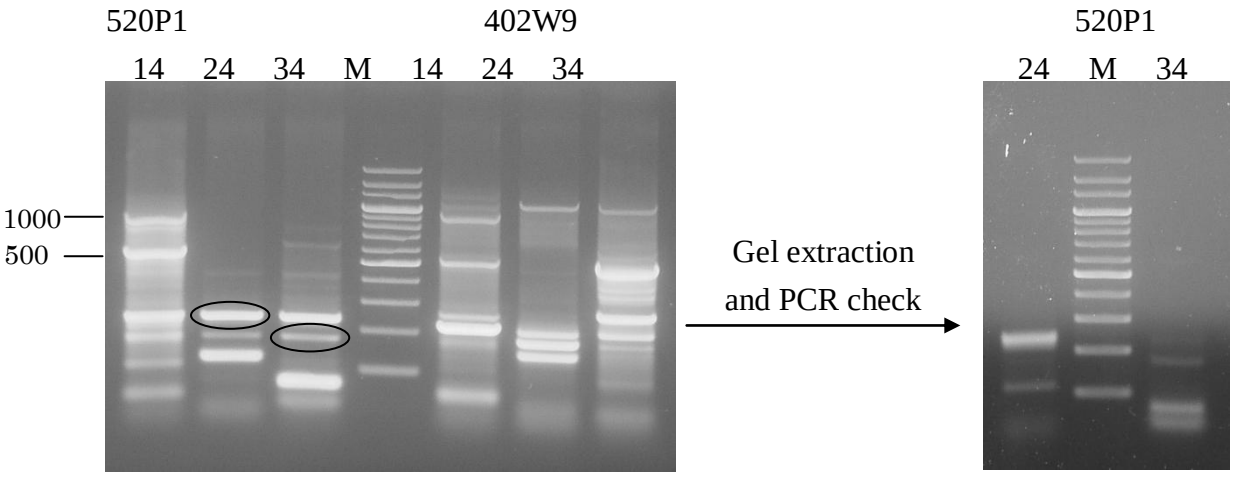


Fig.2.5 The analysis of degenerate PCR products on agarose gel 1.8%

M: DNA marker (200 bp)
14: region 1 →4 (expected size: 411 bp); 24: region 2 →4 (expected size: 246 bp)
34: region 3 →4 (expected size: 183 bp)

The results in **Fig. 2.5.A** showed that there were two positive DNA bands in lanes 24 and 34 of samples from strain 520P1. The sizes of these band were approximately 246 bp and 183 bp, equivalent to the expected sizes of nucleotide sequences in theory between regions 2 to 4 and 3 to 4 of *luxI* gene homologue in strain 520P1. No positive DNA bands with expected size appeared on the lanes of samples from positive control *Vibrio* sp. 402W9.

Then, two putative DNA fragments of lanes 24 and 34 from strain 520P1 were extracted by gel extraction kit and checked again by PCR. The analysis of the PCR products on agarose gel electrophoresis showed that these DNA bands are not specific and contain several other DNA fragments (**Fig.2.5.B**).

2.4. Discussion

Although the LuxI family proteins of *Pseudoalteromonas* strains share several conserved amino acid regions, the result of degenerate PCR indicated that the partial regions of *luxI* gene homologue in *Pseudoalteromonas* sp. 520P1 could not be amplified using degenerated primers. The reason for this failure might come from the low homology of LuxIs at nucleotide level.

Therefore, to identify and characterize the *luxI* gene homologue in strain 520P1, it is essential to sequence the whole genome of *Pseudoalteromonas* sp. 520P1.

CHAPTER III

Genome Sequencing of *Pseudoalteromonas* sp. 520P1 No. 412

3.1 Introduction

In a previous study, Wang et al. (2008) reported that the production of violacein by strain 520P1 was regulated by quorum-sensing mechanisms using an *N*-acylhomoserine lactone (AHL). In *Vibrio fischeri*, two essential components in quorum sensing-regulated bioluminescence, namely AHL synthase (LuxI) and AHL receptor protein (LuxR), and their genes (*luxI*/*luxR*) have been revealed. However, homologous genes for *luxI* and *luxR* in strain 520P1 have not been reported so far. Identification of these genes is pivotal to understand regulatory mechanisms of quorum sensing and the nature of AHL(s) involved in violacein production.

Due to the low homology of LuxIs at nucleotide level, we could not to identify the *luxI* gene homologue in strain 520P1 by degenerate PCR. Therefore, we sequenced the whole genome of *Pseudoalteromonas* strain 520P1 No.412 (NBRC 107704) to identify the *luxI* and *luxR* genes.

3.2. Materials and Methods

3.2.1. GenBank/DDBJ accession numbers

The accession numbers of protein sequences in DDBJ/EMBL/GenBank used in this study were P54656 (EsaI, *Pantoea stewartii* subsp. *stewartii*), ABG38892 (LuxI, *Pseudoalteromonas atlantica* T6c), ERG18256 (LuxI, *Pseudoalteromonas citrea* DSM8771), WP_046355763 (LuxI, *Pseudoalteromonas luteoviolacea*), YP_206882 (LuxI, *Vibrio fischeri* ES114), WP_005423459 (LuxI, *Aliivibrio fischeri*), ABG38890 (LuxR, *P. atlantica* T6c), ERG18412 (LuxR, *P. citrea* DSM8771), CCQ09547 (LuxR, *P. luteoviolacea* B, ATCC 29581), P35327 (LuxR, *Vibrio fischeri* ES114) and ACH63788

(LuxR, *Vibrio fischeri* MJ11).

3.2.2. Sequencing of the whole genome of strain 520P1 No. 412

Genomic DNA of strain 520P1 No. 412 was purified using a QIAGEN Genomic DNA kit with a Genome-tip 100/G column (Qiagen KK, Tokyo, Japan). The genome was sequenced on an Illumina Hiseq 2000 system by Macrogen Japan (Tokyo, Japan). A total of 5,740,346 reads were assembled using SOAPdenovo Assembly into 67 scaffolds with an N_{50} length of 136,339 bp (Luo et al. 2012). The assembled draft genome sequence was approximately 5.25 Mb long with a total coverage of 110-fold and a G+C content of 34.96%. A total of 4,899 protein-coding regions and 99 RNA-coding sequences were detected using the GLIMMER system and RAST Server respectively (Delcher et al. 2007; Overbeek et al. 2014).

3.2.3. Phylogenetic analysis

Homologous sequences were multiply aligned using Clustal X2 (Larkin et al. 2007). Conserved regions and gaps were determined through alignment of amino acid sequences of LuxI and LuxR families. Phylogenetic trees were constructed based on the neighbor-joining method (Saitou and Nei, 1987) using Clustal X2.

3.3. Results

3.3.1. *luxI* and *luxR* gene homologues in the genome of strain 520P1

Genome analysis and annotation of ORFs in 67 scaffolds of *Pseudoalteromonas* sp. 520P1 No. 412 revealed the presence of at least one *luxI* homologue and five *luxR* homologues in the genome of strain 520P1 No. 412.

Table 3.1 BLAST search of *luxI* and *luxR* homologues

	luxI homologues	luxR homologues
Scaffold 1		orf00001 (*palR1*)

Scaffold 15		orf00056 (*palR2*)
Scaffold 19	orf00203 (*palI*)	orf00201 (*palR3*)
Scaffold 51		orf00116 (*palR4*)
Scaffold 56		orf00026 (*palR5*)

The results of genomic analysis also indicated that a pair of *luxI/luxR* homologues was located on the scaffold 19 of the whole genome of strain 520P1 No. 412 (GenBank: DF820557.1). The distance between *luxI* and *luxR* in scaffold 19 is 1021 bp.

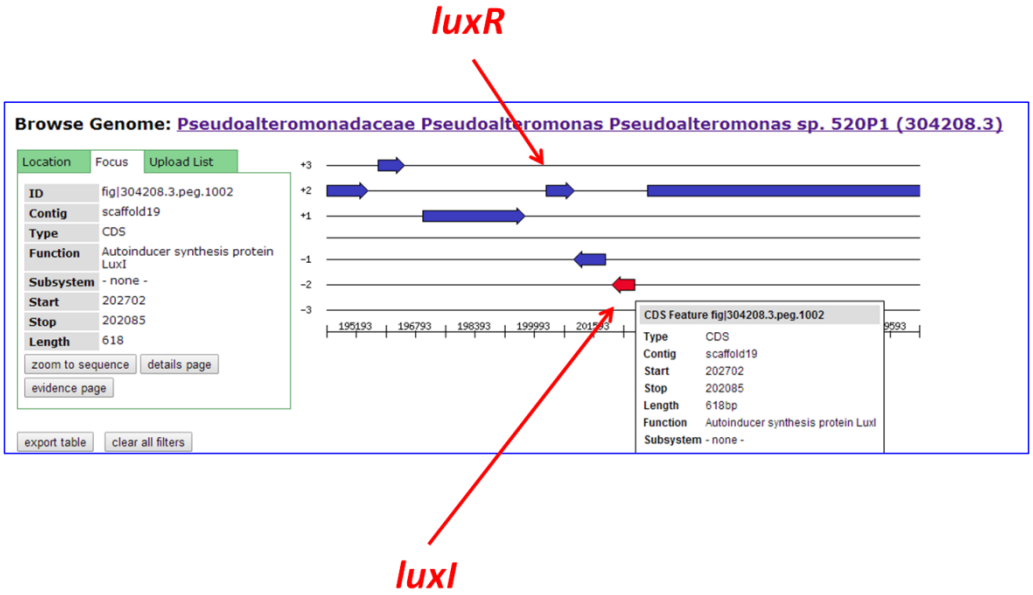


Fig.3.1 Location of luxI/luxR genes on scaffold 19

The genome sequence of strain 520P1 No. 412 has been deposited in DDBJ/EMBL/GenBank under the accession number BBIN01000000 (Dang et al. 2014).

The *luxI* and five *luxR* gene homologues identified in strain 520P1 were named *palI* and *palR1* to *palR5* after the name of genus *Pseudoalteromonas*. The nucleotide sequences of *palI* and five *palR* genes were deposited in DDBJ under accession numbers LC081974, LC081975, LC081976, LC081977, LC081978 and LC081979, respectively.

3.3.2. Homology comparison of LuxI family proteins in strain 520P1 and reference strains

Among AHL synthase proteins of *Pseudoalteromonas* strains, three LuxI family proteins from *P. atlantica* T6c, *P. citrea* DSM8771 and *P. luteoviolacea* have been determined.

Another LuxI homologue, EsaI (PDB code: 1KZF), produces a 3-oxo-hexanoyl-homoserine lactone, which contributes to the quorum sensing regulation of pathogenicity in *Pantoea stewartii* subsp. *stewartii* (Beck von Bodman and Farrand 1995). The structure of EsaI was determined using X-ray crystallography and is refined at a resolution of 1.8 Å. Six amino acid regions (region I – VI) in EsaI have been determined to play an important role in the activity of AHL synthase (Watson et al. 2002; Chakrabarti and Sowdhamini 2003).

Protein sequences of AHL synthase superfamily (EsaI, LuxI homologues from *P. atlantica*, *P. citrea* and *P.luteoviolacea*) were employed for comparison of the protein sequence deduced from the nucleotide sequence of *palI* (*luxI* homologue) of scaffold 19.

The result of amino acid alignment showed that the sequences of LuxI family shares the six motifs of short amino acid sequences that were found in the structure of EsaI (**Fig. 3.2**).

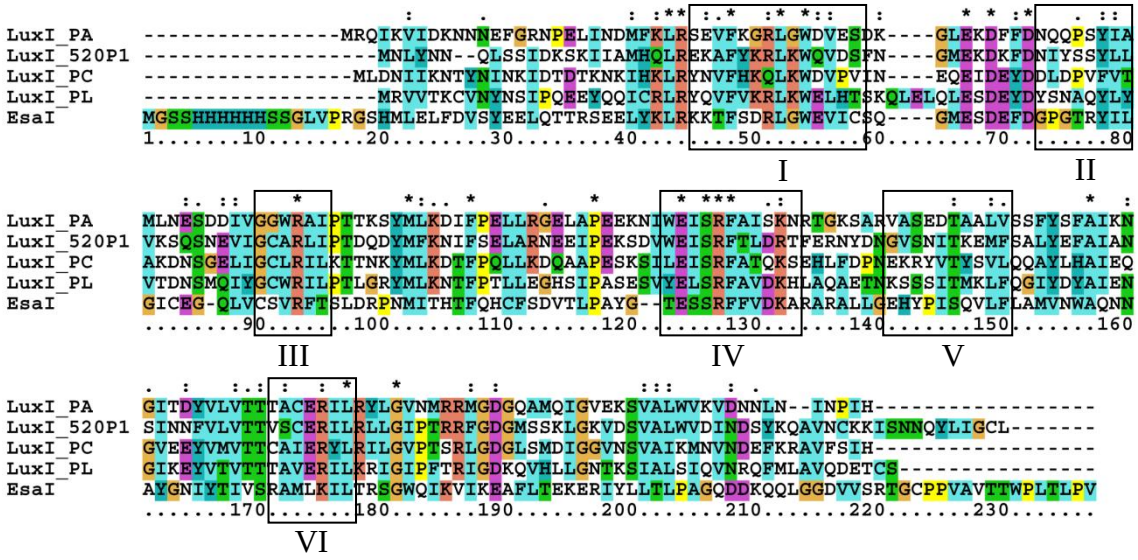


Fig.3.2 Alignment of amino acid sequences of LuxI family.

PA, *P. atlantica* T6c; PC, *P. citrea* DSM8771; PL, *P. luteoviolacea*; EsaI, LuxI homologue from *Pantoea stewartii*;
520P1, PalI of scaffold 19 from *Pseudoalteromonas* sp. 520P1;

The alignment between amino acid sequences of LuxI family revealed that the lowest score of 19 % homology between the LuxI homologue of *P. citrea* and EsaI. The highest score of 45 % homology was obtained between the PalI protein in strain 520P1 and the LuxI homologue in *P. atlantica* (**Fig.3.3**).

	PA	PC	PL	520P1	Esal
PA	**100**	41	37	**45**	21
PC	41	**100**	39	38	19
PL	37	39	**100**	37	28
520P1	**45**	38	37	**100**	24
Esal	21	19	28	24	**100**

Fig.3.3 Matrix of protein homology comparison of LuxI family.

PA, *P. atlantica* T6c; PC, *P. citrea* DSM8771; PL, *P. luteoviolacea*; EsaI, LuxI homologue from *Pantoea stewartii*; 520P1, PalI of scaffold 19 from *Pseudoalteromonas* sp. 520P1

Phylogenetic analysis of the amino acid sequences of LuxI family proteins was performed. As shown in Fig. 1B, the phylogenetic tree indicated the relationship between strain 520P1 and *P. atlantica* (**Fig. 3.4**).

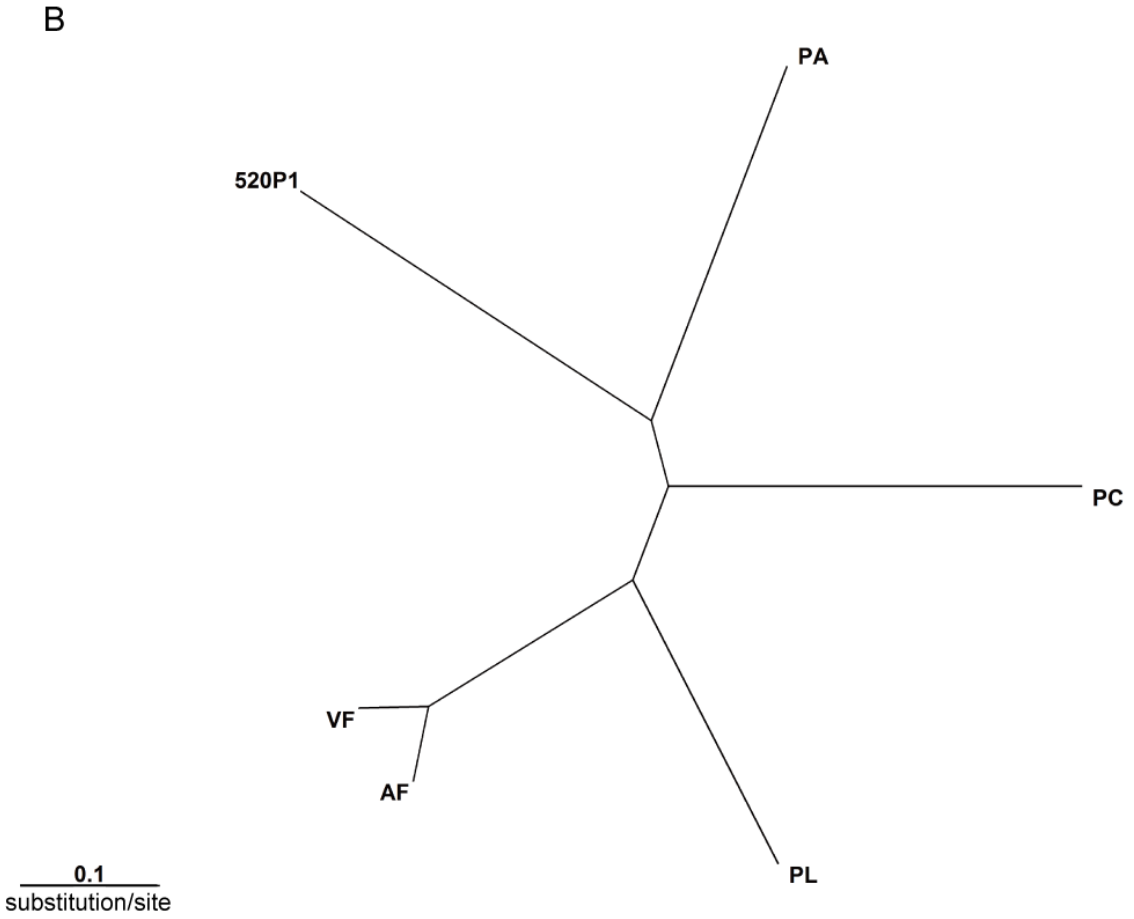


Fig.3.4 Phylogenetic analysis of LuxI family.

PA, *P. atlantica* T6c; PC, *P. citrea* DSM8771; PL, *P. luteoviolacea*; 520P1, PalI of scaffold 19 from *Pseudoalteromonas* sp. 520P1; VF, *V. fischeri* ES114; AF: *A. fischeri*.

3.3.3. Homology comparison of LuxR family proteins in strain 520P1 and reference strains

As described in previous part, the results of genome analysis and annotation revealed the presence of five *luxR* gene homologues (*palR1* to *palR5*) in the genome of strain 520P1 No. 412. The genes *palR1*, *palR2*, *palR3*, *palR4*, *palR5* were located on scaffolds 1, 15, 19, 51, 56, respectively. The amino acid sequences of PalR proteins were

also deduced from their nucleotide sequences.

In the previous structural studies of LuxR proteins, Fuqua et al. (1994) indicated that the C-terminal domain is sufficient for DNA binding and transcriptional activation. They also found that the sequence from residues 190 to 200 contains a helix-turn-helix motif with similarity to the DNA-binding regions of several other transcription factors. In addition, other studies showed that a LuxR fragment harboring only the C-terminal region is still active and unaffected by AHLs (Hanzelka et al. 1997). This indicates that the C-terminal domain includes essential sites for LuxR-DNA and LuxR-RNA polymerase interactions (Choi and Greenberg 1991; Finney et al. 2002; Waters and Bassler 2005).

Alignment of amino acid sequences of PalR proteins of strain 520P1 and LuxR protein homologues in reference *Pseudoalteromonas* strains was conducted. The result of homology comparison showed that they generally share low homology of amino acid sequences. However, several conserved positions could be observed in the C-terminal regions of all sequences (**Fig. 3.5**).

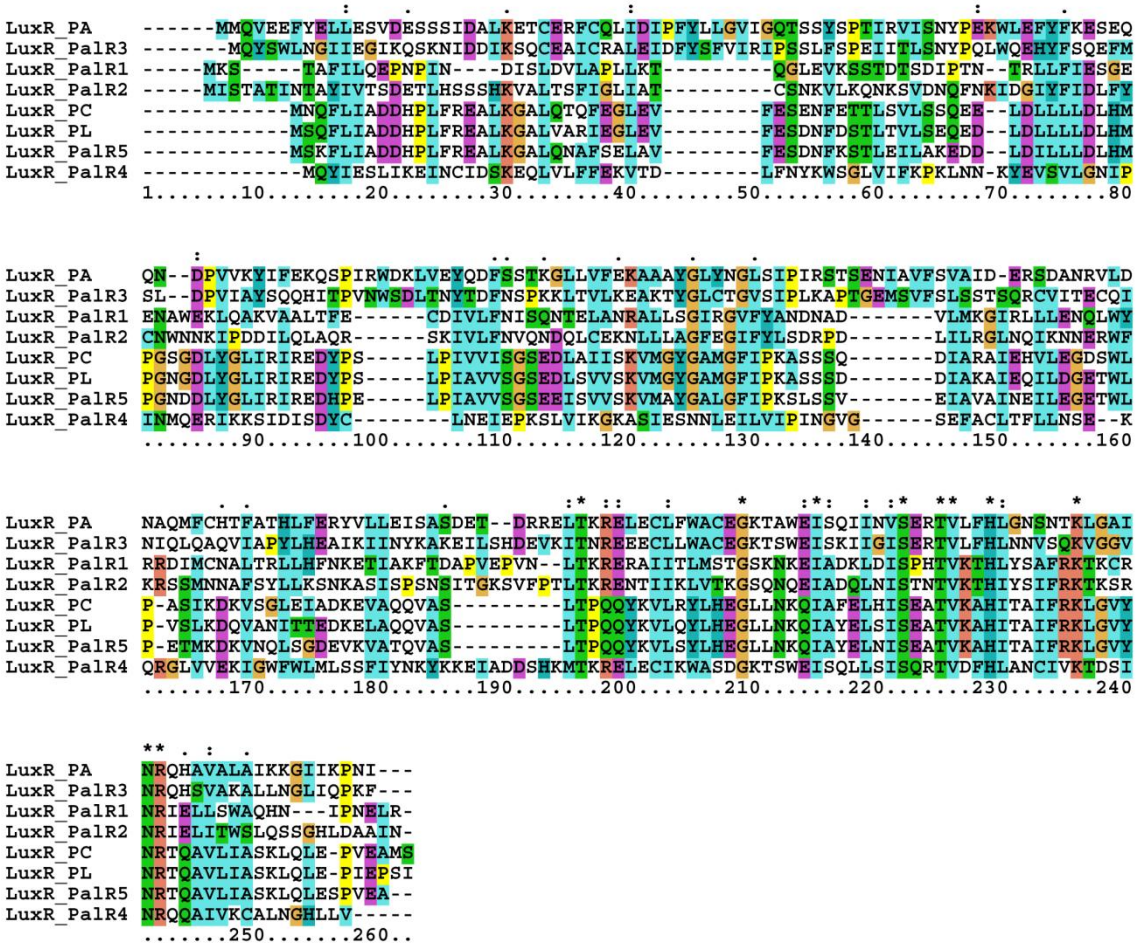


Fig.3.5 Alignment of amino acid sequences of LuxI family.

PA, *P. atlantica* T6c; PC, *P. citrea* DSM8771;

PL, *P. luteoviolacea*; PalR1 – PalR5, LuxR homologues from *Pseudoalteromonas* sp. 520P1.

The alignment also demonstated that PalR5 and LuxR homologues from *P. citrea* DSM8771 and *P. luteoviolacea* share 74 % and 77 % of homology, respectively. In addition, LuxR homologues from *P. citrea* DSM8771 and *P.luteoviolacea* showed close homology of 82 % (**Fig.3.6**)

	PalR1	PalR2	PalR3	PalR4	PalR5	PA	PC	PL
PalR1	**100**	44	41	22	27	21	25	25
PalR2	44	**100**	17	24	30	36	28	33
PalR3	41	17	**100**	31	30	40	31	22
PalR4	22	24	31	**100**	16	35	20	18
PalR5	27	30	30	16	**100**	34	**74**	**77**
PA	21	36	40	35	34	**100**	27	34
PC	25	28	31	20	**74**	27	**100**	**82**
PL	25	33	22	18	**77**	34	**82**	**100**

Fig.3.6 Phylogenetic analysis of LuxR family.

PA, *P. atlantica* T6c; PC, *P. citrea* DSM8771; PL, *P. luteoviolacea*; PalR1 – PalR5, LuxR homologues from *Pseudoalteromonas* sp. 520P1.

Phylogenetic analysis of the amino acid sequences of LuxR family was performed. The phylogenetic tree indicated the close relationship between LuxR family protein of three strains, including PalR5 of strain 520P1 No. 412, and LuxR homologues of *P. citrea* DSM8771 and *P. luteoviolacea* (**Fig. 3.7**).

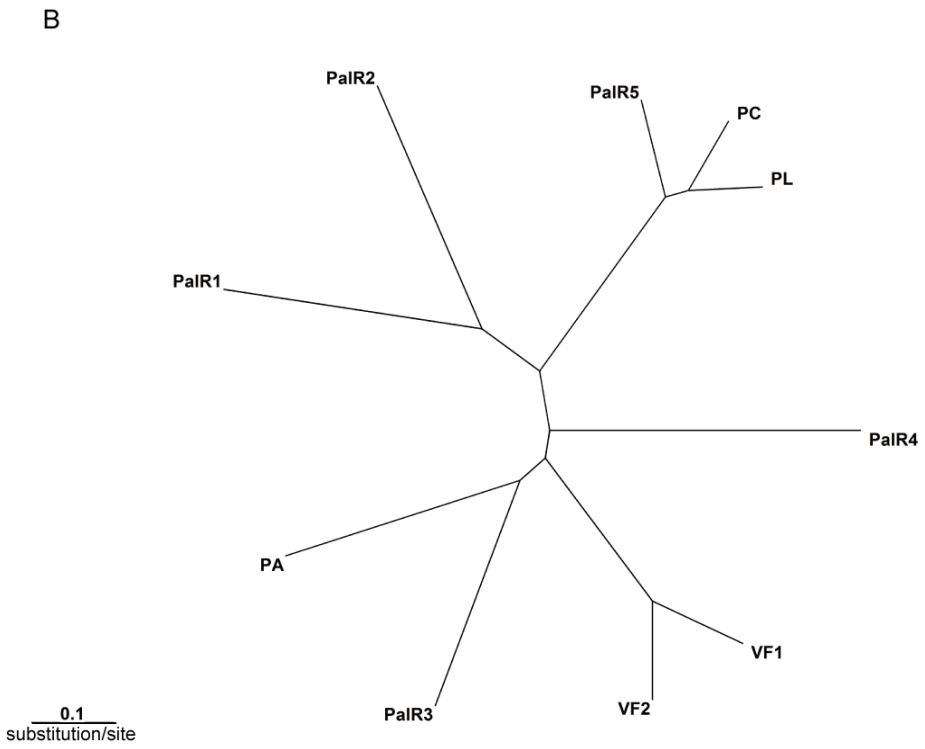


Fig.3.7 Phylogenetic analysis of LuxR family.

PA, *P. atlantica* T6c; PC, *P. citrea* DSM8771; PL, *P. luteoviolacea*; VF1, *V. fischeri* ES114; VF2, *V. fischeri* MJ11; PalR1 – PalR5, LuxR homologues from *Pseudoalteromonas* sp. 520P1.

3.4. Discussion

In the present study, we characterized one LuxI and five LuxR homologues in *Pseudoalteromonas* sp. 520P1. These homologues were named PalI and PalR1 to PalR5, respectively. The properties of PalI and five PalRs provide the basis for understanding the components of *N*-acylhomoserine lactone-dependent quorum sensing system in this marine bacterium.

We compared and analyzed the amino acid sequence of PalI with LuxI homologues from different bacterial species. Based on the structure of EsaI, an AHL synthase well-defined using X-ray crystallography, we identified six amino acid regions that share significantly conserved homology in all aligned sequences. The results suggest that PalI is an AHL synthase and has a close relationship with LuxI from *P. atlantica* T6c

among LuxI homologues in *Pseudoalteromonas*.

In the most common quorum sensing system in Gram-negative bacteria, LuxI and LuxR generally exist in cognate pairs. LuxI homologue in the pair produces a specific AHL required to activate the corresponding LuxR protein. In most cases, cognate *luxI* and *luxR* genes are located adjacently to each other (Gray and Garey 2001; Lerat and Moran 2004). However, the sequencing of many bacterial genomes has revealed that many proteobacteria only contain LuxRs that do not have a cognate LuxI protein associated with them. These LuxRs have been called orphans and more recently solos. LuxR solos are found to be widespread in proteobacterial species that possess an *N*-acylhomoserine lactone-dependent quorum sensing system as well as in species that do not. Genomic analysis of 265 bacterial strains revealed that 68 strains have at least one LuxI and one LuxR homologue (Case et al. 2008). Notably, 45 of these 68 strains have more than one LuxR solo. This finding indicates that these bacterial genomes probably contain one or more LuxR solos. In addition, SdiA, the LuxR solo of *E. coli* and *Salmonella enteric*, has beenproved to have an ability to respond to exogenous AHLs synthesized by other bacteria (Michael et al. 2001).

In recent study, it has been reported that LuxR-family proteins usually differ in their length and are made up of two domains; the N-terminus containing an AHL-binding domain and the C-terminus possessing a helix-turn-helix DNA-binding domain. Remarkably, LuxRs share low similarities (18-25%) even if they respond to structurally similar AHLs (Vannini et al. 2002; Subramoni and Venturi 2009; Brameyer et al. 2014).

In this study, we identified five LuxR homologues (PalR1 to PalR5) in the genome of the strain 520P1, in which one pair of PalI and PalR3 are located adjacently on the scaffold 19. When compared with LuxRs from other *Pseudoalteromonas* species, all

sequences share low homology in the AHL-binding domain (N-terminus) but retain some conserved amino acid sequences in the DNA-binding domain (C-terminus). The divergence of N-terminal regions in LuxR homologues of strain 520P1 is consistent with the general characteristics of LuxR family proteins from different species, including *Pseudoalteromonas* species.

Among LuxR homologues of strain 520P1, we found PalR5 shares 74 % and 77 % homology with LuxRs from *P. citrea* DSM8771 and *P. luteoviolacea*, respectively. On the other hand, phylogenetic analysis showed that other LuxR homologues of strain 520P1, PalR1, PalR2, PalR3 and PalR4, were not in the same branch with PalR5 and phylogenetically separated from each other. The reason for this could be explained when we assume that *palR* genes have been acquired from different sources during the evolutionary process of *Pseudoalteromonas* sp. 520P1 probably by horizontal gene transfer. Since the cognate LuxI homologues of these PalRs except PalR3 have not been identified, it is most likely that they are LuxR solos. In the strain 520P1, these LuxR solos might respond to endogenous and exogenous signals produced by neighboring bacteria and then control a set of genes that overlaps the set regulated by PalR3.

Although the precise manner by which each of PalRs in the strain 520P1 is regulated has yet to be elucidated, these findings contribute to clarify the components of *N*-acylhomoserine lactone-dependent quorum sensing system in the marine bacterium *Pseudoalteromonas* sp. 520P1.

CHAPTER IV

Cloning and expression of the *PalI* gene from *Pseudoalteromonas* sp. 520P1 in *E. coli*

4.1 Introduction

In the previous work, the *palI* gene coding for AHL synthase in *Pseudoalteromonas* 520P1 was identified to exist in pair with the palR3 in the scaffold 19. The production of violacein by strain 520P1 has been reported to be regulated by quorum-sensing mechanisms using an *N*-acylhomoserine lactone (AHL) (Wang et al. 2008).

To examine whether the *palI* gene product indeed has AHL synthase activity, we tried to clone this gene and express in *E. coli* BL21 (DE3).

4.2. Materials and Methods

4.2.1. Bacterial strains, growth conditions and plasmids

Genomic DNA of *Pseudoalteromonas* 520P1 No. 412 (NBRC 107704) was used to sequence the whole genome and the cloning of the *palI* gene. *Agrobacterium tumefaciens* NTL4 (pZLR4) was used as a reporter strain to detect AHLs according to the method of Wang et al. (2008).

Plasmid pUC18 and *Escherichia coli* DH5α were purchased from Takara Bio (Otsu, Japan). Plasmid pET28a (+) and *E. coli* BL21 (DE3) were purchased from Novagen (Darmstadt, Germany). KOD-plus DNA polymerase for PCR amplification was purchased from Toyoba (Osaka, Japan). Restriction enzymes, a DNA ligation kit (Mighty Mix), X-gal and IPTG were purchased from Takara Bio. Ampicillin (100 μg/ml) and kanamycin (50 μg/ml) were used for the selection of recombinant *E. coli*.

4.2.2. Cloning of *palI* gene of strain 520P1 No. 412

Based on the sequence of the *palI* gene in strain 520P1, two pairs of primers were designed and used to amplify the DNA region containing this gene by nested PCR. In the first PCR, the genomic DNA of *Pseudoalteromonas* sp. 520P1 No. 412 was used as a template. The DNA fragment containing upper and lower regions of the *palI* gene was amplified using the forward primer S19O203-ex-Fw (5'-TTGGGTACAAGATAGTTTGTTATA-3') and the reverse primer S19O203-ex-Rv (5'-TAGCGATAAAGAATGCTTAATATA-3'). Then, PCR product was purified and used as a template for the second PCR. The reaction was carried out with the forward primer NdeI-Fw-19203 (5'-CGC**CATATG**ATGAATCTCTATAAC-3') and the reverse primer HindIII-Rv-19203 (5'-GCG**AAGCTT**TAAACAACCAATAAG-3'). The product of the second PCR containing restriction enzyme sites on both ends of the *palI* gene was named NdeI-palI-HindIII. The gene NdeI-palI-HindIII was digested with *Nde*I and *Hind*III and cloned into the multiple cloning sites of pUC18 plasmid to create the cloning plasmid pUC18-NdeI-palI-HindIII. Then, the recombinant plasmid was transformed into competent *E. coli* DH5α by heat shock method. All PCR were performed with fidelity KOD-Plus DNA polymerase. PCR products were purified using a PCR purification kit or a gel extraction kit (Qiagen KK, Tokyo, Japan).

4.2.3. Construction of the expression plasmid for *E. coli*

To examine the AHL production by the cloned *palI* gene, a recombinant pET28a(+) expression plasmid containing the *palI* gene was constructed. PCR-amplified NdeI-palI-HindIII from pUC18-NdeI-palI-HindIII was digested by *Nde*I and *Hind*III, and cloned into the cloning sites of the pET28a(+) vector to produce a recombinant plasmid pET28a-NdeI-palI-HindIII. The recombinant vector was

constructed in *E. coli* DH5α and then transformed into *E. coli* BL21 (DE3) for expression. DNA sequencing of the insert in the pET28a(+) vector was performed using the dye-terminator method by Bio Matrix Research (Nagareyama, Japan).

4.2.4. Expression of *palI* gene

A single colony of *E. coli* BL21 (DE3) harboring the recombinant plasmid pET28a-NdeI-palI-HindIII was inoculated into 200-ml flasks containing 50 ml of fresh LB medium. The flasks were incubated on a shaker at 37°C until the optical density at 600 nm (OD_{600}) of the cultures reached 0.9- 1.0. Then, 0.5 mM IPTG was added to the flasks and the culture for AHL production was carried out at 20°C with shaking (180 rpm) for 24 h. AHLs were extracted from the culture supernatant and detected as described below.

4.2.5. Extraction of AHLs

The culture supernatants (100 ml each) were harvested by centrifugation at 8,000 rpm (11,800 × g) for 30 min. Then, AHLs in the supernatants were extracted with the same volume of ethyl acetate in a separating funnel. The ethyl acetate phase was recovered and evaporated to dryness. The dried samples were dissolved in 1.0 ml each of ethanol. The extracts were stored at −20°C until use. AHL was also extracted from strain 520P1 No. 412 after 72 h of static culture as described above.

4.2.6. Plate assay of AHLs

Agrobacterium tumefaciens NTL4 (pZLR4) was cultured in AB medium containing 0.2 % glucose and 1 μg/ml gentamicin on a shaker at 200 rpm at 28ºC. An AB minimal agar plate containing 0.2 % glucose and 40 mg/ml of X-gal was overlaid with 3 ml of molten soft agar containing 40 mg/ml of X-Gal and 200 μl of overnight culture of *A. tumefaciens* NTL4 (pZLR4). After the soft agar was solidified, paper discs (8 mm in

diameter) filled with 20 µl each of AHL extracts were laid on the surface of the plate. Following incubation for 24–48 h, the plate was examined for zones of blue pigmentation (Fuqua and Winans 1994; Wang et al.2008).

4.2.7. TLC assay of AHLs

Reverse-phase (RP)–TLC plates were used to analyze AHLs. The AHL extracts were spotted on a TLC plate (Silica Gel 60 RP-18F_{254}S Merck, Darmstadt, Germany), and the plate was developed in a solvent system of methanol:water (60:40, v/v). After chromatography, molten soft agar containing X-gal and *A. tumefaciens* NTL4 (pZLR4) was poured over the surface of the TLC plate as describe above, and the TLC plate was incubated at 28°C for 24–48 h in a closed plastic container (Shaw PD et al. 1997; Wang et al. 2008).

4.3. Results

4.3.1. Cloning of *palI* gene from strain 520P1

Genomic DNA isolated from 520P1 No. 412 was used as a template for the amplification of *palI* by two steps of PCR with specific primers. The size of the product of the first PCR including upper and lower regions of the *palI* is 738 bp. Then, this amplified fragment was used as a template for the second PCR in which restriction enzyme sites were added to the both ends of the PCR product. The size of the product of the second PCR is 636 bp in which the full length of the *palI* is 618 bp. After that, purified DNA fragment of the second PCR was cloned into pUC18 plasmid and subcloned into pET28a(+) plasmid for protein expression. The results of DNA sequencing of the inserted DNA showed that the determined nucleotide sequence was identical to the designed one and correctly orientated in the expression plasmid (Fig. 4.1).

Fig.4.1 Schematic diagram of the *palI* gene in pET28a(+) vector.

To examine whether the protein encoded by the cloned *palI* gene has AHL synthase activity, the expression of *palI* gene was conducted in an *E. coli* strain which does not produce AHLs.

4.3.2. Expression of cloned *palI* gene in *E. coli* and detection of AHL activity

AHLs as signaling molecules play an important role in the quorum sensing mechanism. Therefore, various bacterial biosensors of AHLs, which do not produce AHLs by themselves but respond to exogenous AHLs, have been developed to detect AHLs. In this study, AHLs were extracted from culture supernatant and detected by plate assay of AHL using *Agrobacterium tumefaciens* NTL4 (pZLR4) as an AHL-reporter strain.

A. tumefaciens NTL4 (pZLR4) contains pZLR4 plasmid which confers resistance to gentamicin. pZLR4 includes a *traG*/*lacZ* fusion gene and *traR*, a LuxR homologue in *A. tumefacience*. When AHL is added to strain NTL4 (pZLR4), it binds the TraR protein, and the complex of Tra/AHL promotes transcription of the *traG/lacZ* fusion gene. Thus, β-galactosidase activity expressed by *traG/lacZ* fusion gene can be used as an indicator of AHLs (Fuqua and Winans 1994; Miller and Bassler 2001; Farrand et al. 2002; Steindler and Venturi 2007). In quantification plate assay, the soft agar containing overnight culture of *A. tumefaciens* NTL4 (pZLR4) was overlaid on the AB minimal agar plate. Then, the paper discs filled with AHL extracts were laid on the surface of the plate.

Fig.4.2 shows the detection of AHLs extracted from the culture of recombinant *E. coli* after 24 h of culture in the presence of 0.5 mM IPTG. An AHL standard (3-oxo-C8-HSL) was used as a positive control. A signal of AHL from the recombinant *E. coli* was obtained as a blue halo on the indicator plate, while no activity was observed when the extract of control *E. coli* harboring pET28a(+) without *palI* gene was used.

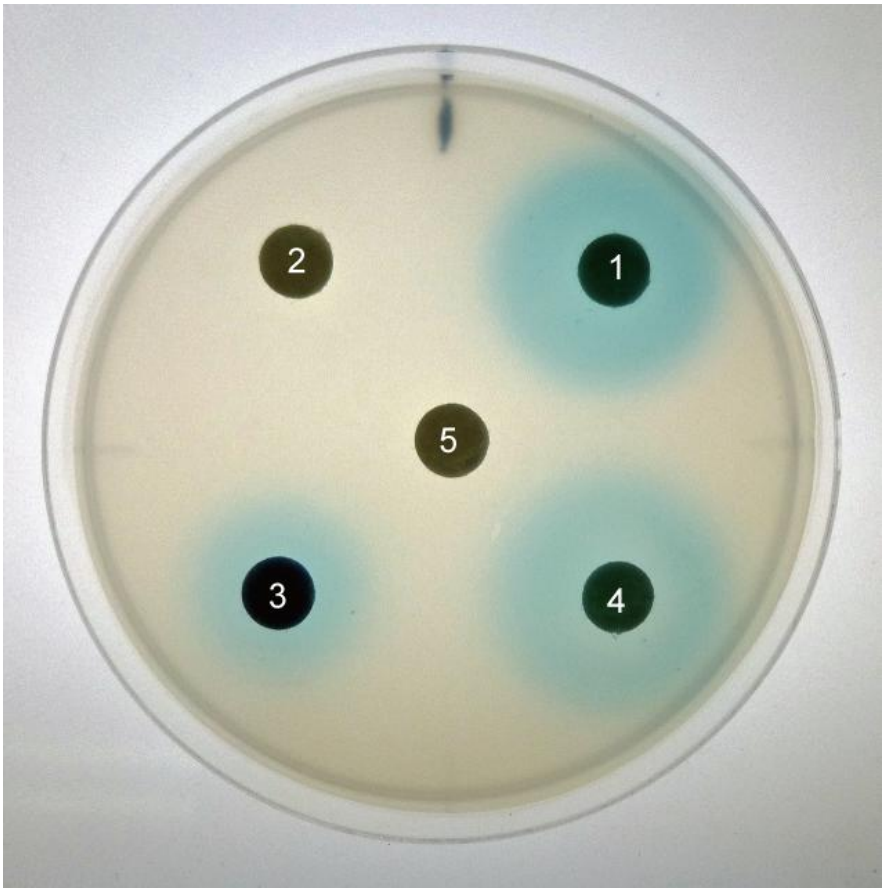


Fig.4.2 Bioassay of AHL activity in the culture supernatant of recombinant *E. coli*. Spot 1, AHLs produced by the recombinant *E. coli* (10 µl extract, 24 h of culture); spot 2, extract from *E. coli* containing pET28a without the insert (negative control, 10 µl extract, 24 h of culture); spot 3, AHLs produced by strain 520P1 No. 412 (100 µl extract, 72 h of culture); spot 4, standard 3-oxo-C8 HSL (0.1 ng); spot 5, ethanol (negative control, 100 µl)

4.3.3. Analysis of AHLs by TLC

In the previous studies, based on the data of TLC and LC-MS, Wang et al. (2008)

reported that 3-oxo-C8-HSL and C14-HSL were present in the culture supernatant of strain 520P1. This data suggests that the strain 520P1 can produce several types of AHL autoinducers.

As shown in Fig.4.3, the analysis of the AHL extracts by TLC plate demonstrated that two spots appeared on the lane of the extract from *E. coli* (lane 3), in which the two spots were similar in mobility to those of AHLs from strain 520P1 No. 412 (lane 1). This result indicates that the *palI* gene of scaffold 19 in strain 520P1 is able to express at least two types of AHLs in *E. coli*.

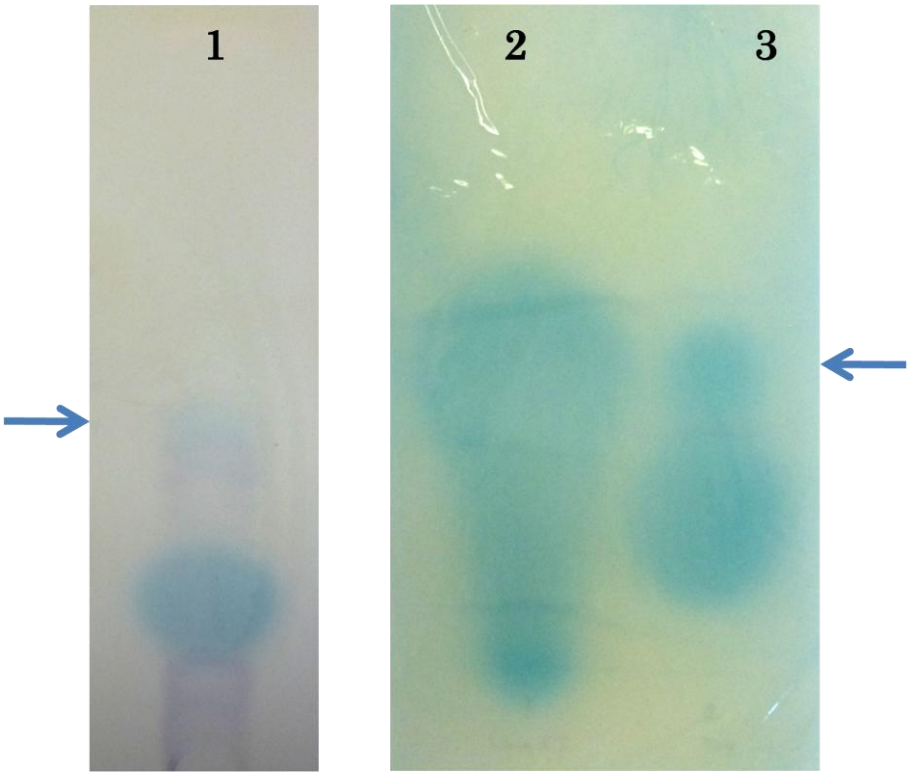


Fig.4.3 Analysis of AHLs on a reversed phase TLC plate.

lane 1, AHLs extracts from strain 520P1 No. 412 (60 µl extract); lane, 2, standard 3-oxo-C8-HSL (0.1 ng); lane 3, AHL extract from the recombinant *E. coli* (10 µl extract)

4.4. Discussion

We successfully cloned the *palI* gene and expressed in *E. coli* BL21 (DE3). Using

AHL bioassay coupled to TLC plate, two types of AHL were detected in the extract of the expression culture of recombinant *E. coli* harboring *palI* gene. Two spots of AHL on the TLC plate were similar in mobility to those of strain 520P1 No. 412.

In the previous study, Wang et al. (2008) reported two spots of AHL on a TLC plate, corresponding to 3-oxo-C8-HSL and C14-HSL. From this observation, it was considered that two LuxI homologues could be present and each LuxI homologue produced corresponding AHL molecule. However, bioassay of AHL on a TLC plate in this study demonstrated that PalI can produce in *E. coli* two types of AHL similar to those in strain 520P1. Therefore, possible involvement of PalI in the regulation of violacein synthesis in strain 520P1 is suggested if we can assume that PalI is the only AHL synthase in strain 520P1.

CHAPTER V

Genome sequencing of a variant of *Pseudoalteromonas* sp. 520P1

5.1 Introduction

In the previous study, strain 520P1 was found to produce violacein under static culture conditions but hardly to produce violacein with agitated culture conditions. From a culture flask of strain 520P1, Zhang (2010) isolated a variant strain with an ability of highly stable production under agitated culture conditions. This variant strain probably generated by natural mutation was named strain 520P1 No. 423, while a strain with original phenotypic features in the same culture flask was called strain 520P1 No. 412. The variant strain 520P1 No. 423 produced much amount of AHLs than strain 520P1 No. 412. Therefore, it was considered that overproduction of AHLs by the strain 520P1 No. 423 led to the production of violacein under agitated culture conditions.

These properties of the variant strain in producing violacein have provided an advantage to use strain 520P1 No. 423 for a large scale production of violacein.

In the previous study, it was also hypothesized that a mutation in the upstream promotor region of violacein gene cluster in strain 520P1 No. 423 led to the production of violacein under agitated culture conditions. However, nucleotide sequencing showed that the upstream sequence of violacein gene cluster of strain 520P1 No. 423 was identical to that of the original strain 520P1 No. 412. Therefore, no mutation occurred in the upstream region of violacein gene cluster ofstrain 520P1 No. 423.

To clarify the difference of *palI* and *palR* genes between strain 520P1 No. 412 and No. 423, we performed the genomic sequencing of strain 520P1 No. 423.

5.2. Materials and Methods

Genomic DNA of strain 520P1 No. 423 was purified using a QIAGEN Genomic DNA kit with a Genome-tip 100/G column (Qiagen KK, Tokyo, Japan). The genome was sequenced on a PacBio RSII system by Macrogen Japan (Tokyo, Japan).

5.3. Results

5.3.1 Whole genome analysis of strain 520P1 No. 423

A total of 111,839 reads were assembled using HGAP3 into 2 contigs with an N50 length of 13,115 bp. The assembled draft genome sequence was approximately 5.39 Mb long. A total of 4,855 protein-coding regions, 122 tRNA and 34 rRNA genes were detected using the Prokka for the rapid annotation of prokaryotic genomes.

The draft whole genome of strain 520P1 No. 423 has been deposited in DDBJ/EMBL/GenBank under the accession number BBZB01000000.

5.3.2 Comparison with the genome of strain 520P1 No. 412

Homology comparison between partial genomic sequences of strain 520P1 No. 412 and No. 423 was performed (**Appendix 2**). The results showed that no mutation was detected in the nucleotide sequences of the *palI* and its promoter, five *palRs* and their promoters, violacein gene cluster and its promoter. Therefore, unique properties of strain 520P1 No. 423 in violacein production should be described to the gene mutation that is not relevant to *palI*, *palR*s and violacein gene cluster.

5.4. Discussion

With the results of partial genomic analysis, to identify the difference between strain 520P1 No. 412 and No. 423, we need to consider other possible reasons. One of them is the difference in the sequences of AHL-degrading enzymes in strain 520P1 No.

423. In the previous studies, enzymatic degradation of AHL autoinducers has been described (Czajkowski and Jafra 2009). There are two group of AHL-inactivating enzymes have been identified including AHL lactonases which hydrolyze the lactone ring in AHLs and AHL acylases which release a free homoserine lactone and a fatty acid. We suggest that the identification of these enzymes might clarify the ability of producing violacein under agitated culture conditions of strain 520P1 No. 423.

CHAPTER VI

Conclusions

- The whole genome of strain 520P1 No. 412 and No. 423 have been sequenced and deposited in DDBJ/EMBL/GenBank.
- The *palI* gene in scaffold 19 was cloned and expressed in *E. coli*.
- AHL activity was detected in the extract from recombinant *E. coli* by bioassay.
- Five *palR* genes were identified in different scaffolds of strain 520P1 No. 412, in which *palR3* exist as a pair with *palI* in scaffold 19.
- Phylogenetic analysis revealed the high homology between PalR5 in scaffold 56 of strain 520P1 No. 412 with LuxR from *P. citrea* DSM8771 and *P. leuteoviolacea*. This finding contributes to the explanation of the evolutionary history of LuxR family of *Pseudoalteromonas* strains.

REFERENCES

1. August PR, Grossman TH, Minor C, Draper MP, MacNeil IA, Pemberton JM, Call KM, Holt D, Osburne MS (2000) Sequence analysis and functional characterization of the violacein biosynthetic pathway from *Chromobacterium violaceum*. J Mol Microbiol Biotechnol. 2:513-519
2. Beck von Bodman S, Farrand SK (1995) Capsular polysaccharide biosynthesis and pathogenicity in *Erwinia stewartii* require induction by an *N*-acylhomoserine lactone autoinducer. J Bacteriol. 177:5000-8
3. Bottomley MJ, Muraglia E, Bazzo R, Carfì A (2007) Molecular insights into quorum sensing in the human pathogen Pseudomonas aeruginosa from the structure of the virulence regulator LasR bound to its autoinducer. J Biol Chem. 282:13592-600
4. Brameyer S, Kresovic D, Bode HB, Heermann R (2014) LuxR solos in Photorhabdus species. Front Cell Infect Microbiol. 4:166
5. Case RJ, Labbate M, Kjelleberg S (2008) AHL-driven quorum-sensing circuits: their frequency and function among the Proteobacteria. ISME J. 2:345–349
6. Chakrabarti S, Sowdhamini R (2003) Functional sites and evolutionary connections of acylhomoserine lactone synthases. Protein Eng. 16:271-278
7. Choi SH, Greenberg EP (1991) The C-terminal region of the Vibrio fischeri LuxR protein contains an inducer-independent lux gene activating domain. Proc Natl Acad Sci U S A. 88:11115-11119
8. Churchill ME, Chen L (2011) Structural basis of acyl-homoserine lactone-dependent signaling. Chem Rev. 111:68-85
9. Czajkowski R, Jafra S (2009) Quenching of acyl-homoserine lactone-dependent

quorum sensing by enzymatic disruption of signal molecules. Acta Biochim Pol. 56:1-16

10. Dang HT, Yotsumoto K, Enomoto K (2014) Draft genome sequence of violacein-producing marine bacterium *Pseudoalteromonas* sp. 520P1. Genome Announc. 2(6): e01346-14
11. Decho AW, Frey RL, Ferry JL (2011) Chemical challenges to bacterial AHL signaling in the environment. Chem Rev. 111:86-99
12. Delcher AL, Bratke KA, Powers EC, Salzberg SL (2007) Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23:673-679
13. Dong YH, Zhang LH (2005) Quorum sensing and quorum-quenching enzymes. J Microbiol. 43:101-9.
14. Durán M, Menck CF (2001) *Chromobacterium violaceum*: a review of pharmacological and industrial perspectives. Crit Rev Microbiol. 27:201-222
15. Durán M, Ponezi AN, Faljoni-Alario A, Teixeira MFS, Justo GZ, Durán N (2012) Potential applications of violacein: A microbial pigment. Med Chem Res. 21:1524-1532
16. Fast W, Tipton PA (2012). The enzymes of bacterial census and censorship. Trends in Biochemical Sciences. 37:7-14. doi:10.1016/j.tibs.2011.10.001
17. Farrand SK, Qin Y, Oger P (2002) Quorum-sensing system of *Agrobacterium* plasmids: analysis and utility. Methods Enzymol. 358:452–484
18. Finney AH, Blick RJ, Murakami K, Ishihama A, Stevens AM (2002) Role of the C-terminal domain of the alpha subunit of RNA polymerase in LuxR-dependent transcriptional activation of the lux operon during quorum sensing. J Bacteriol.

184:4520-4528

19. Fuqua C, Greenberg EP (2002) Listening in on bacteria: acyl-homoserine lactone signalling. Nat Rev Mol Cell Biol. 3:685-695
20. Fuqua WC, Winans SC (1994) A LuxR-LuxI type regulatory system activates Agrobacterium Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. J Bacteriol. 176:2796–2806
21. Fuqua WC, Winans SC, Greenberg EP (1994) Quorum sensing in bacteria: the LuxR/LuxI family of cell density-responsive transcriptional regulators. J Bacteriol. 176:269–275
22. Galloway WR, Hodgkinson JT, Bowden SD, Welch M, Spring DR (2011) Quorum sensing in Gram-negative bacteria: small-molecule modulation of AHL and AI-2 quorum sensing pathways. Chem Rev. 111:28-67. doi: 10.1021/cr100109t
23. González JE, Marketon MM (2003) Quorum sensing in nitrogen-fixing rhizobia. Microbiol Mol Biol Rev. 67:574-92
24. Gray KM, Garey JR (2001) The evolution of bacterial LuxI and LuxR quorum sensing regulators. Microbiology. 147:2379-2387
25. Hanzelka BL, Greenberg EP (1996) Quorum sensing in *Vibrio fischeri*: evidence that S-adenosyl methionine is the amino acid substrate for autoinducer synthesis. J Bacteriol. 178:5291–5294
26. Hanzelka BL, Stevens AM, Parsek MR, Crone TJ, Greenberg EP (1997) Mutational analysis of the Vibrio fischeri LuxI polypeptide: critical regions of an autoinducer synthase. J Bacteriol. 179:4882-4887
27. Hirano S, Asamizu S, Onaka H, Shiro Y, Nagano S (2008) Crystal structure of VioE, a key player in the construction of the molecular skeleton of violacein. J Biol Chem.

283:6459-6466

28. Hoshino T (2011) Violacein and related tryptophan metabolites produced by *Chromobacterium violaceum*: biosynthetic mechanism and pathway for construction of violacein core. Appl Microbiol Biotechnol. 91:1463-1475
29. Iserte JA, Stephan BI, Goñi SE, Borio CS, Ghiringhelli PD, Lozano ME (2013) Family-specific degenerate primer design: a tool to design consensus degenerated oligonucleotides. Biotechnol Res Int. 2013:383646. doi: 10.1155/2013/383646
30. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics, 23:2947-2948.
31. Lerat E, Moran NA (2004) The evolutionary history of quorum-sensing systems in bacteria. Mol Biol Evol. 21:903-913
32. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu SM, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam TW, Wang J (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience1:18.doi:10.1186/2047-217X-1-18
33. McLean KH, Winson MK, Fish L, Taylor A, Chhabra SR, Camara M, Daykin M, Lamb JH, Swift S, Bycroft BW, Stewart GSAB, Williams P (1997) Quorum sensing and *Chromobacterium violaceum*: exploitation of violacein production and inhibition for the detection of *N*-acylhomoserine lactones. Microbiology. 143:3703–3711
34. Michael B, Smith JN, Swift S, Heffron F, Ahmer BM (2001) SdiA of *Salmonella enterica* is a LuxR homolog that detects mixed microbial communities. J Bacteriol.

183:5733-5742

35. Miller MB, Bassler BL (2001) Quorum sensing in bacteria. Annu Rev Microbiol. 55:165–199
36. Momen AZMR, Hoshino T (2000) Biosynthesis of violacein: Intact incorporation of the tryptophan molecule on the oxindole side, with intramolecular rearrangement of the indole ring on the 5-hydroxyindole side. Biosci Biotechnol Biochem. 64:539-549
37. Morohoshi T, Fukamachi K, Kato M, Kato N, Ikeda T (2010) Regulation of the Violacein biosynthetic gene cluster by acylhomoserine lactone-mediated quorum sensing in *Chromobacterium violaceum* ATCC 12472. Biosci Biotechnol Biochem. 74:2116-2119
38. Ng WL, Bassler BL (2009) Bacterial quorum-sensing network architectures. Annu Rev Genet. 43:197-222
39. Oinuma K, Greenberg EP (2011) Acyl-homoserine lactone binding to and stability of the orphan Pseudomonas aeruginosa quorum-sensing signal receptor QscR. J Bacteriol. 193:421-8. doi: 10.1128/JB.01041-10
40. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R (2014) The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res. 42:D206-D214.doi:10.1093/nar/gkt1226
41. Patankar AV, González JE (2009) Orphan LuxR regulators of quorum sensing. FEMS Microbiol Rev. 33:739–756
42. Pantanella F, Berlutti F, Passariello C, Sarli S, Morea C, Schippa S (2007) Violacein and biofilm production in Janthinobacterium lividum. J Appl Microbiol.

102:992-999

43. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 4:406-425
44. Shaw PD, Ping G, Daly SL, Cha C, Cronan JE Jr, Rinehart KL, Farrand SK (1997) Detecting and characterizing *N*-acylhomoserine lactone signal molecules by thin-layer chromatography. Proc Natl Acad Sci USA. 94:6036-6041
45. Steindler L, Venturi V (2007) Detection of quorum-sensing N-acyl homoserine lactone signal molecules by bacterial biosensors. FEMS Microbiol Lett. 266:1-9
46. Subramoni S, Venturi V (2009) LuxR-family 'solos': bachelor sensors/regulators of signalling molecules. Microbiology. 155:1377-1385
47. Vannini A, Volpari C, Gargioli C, Muraglia E, Cortese R, De Francesco R, Neddermann P, Marco SD (2002) The crystal structure of the quorum sensing protein TraR bound to its autoinducer and target DNA. EMBO J. 21:4393-401
48. Venil CK, Zakaria ZA, Ahmad WA (2013) Bacterial pigments and their applications. Process Biochem. 48:1065-1079
49. Wang Y, Ikawa A, Okaue S, Taniguchi S, Osaka I, Yoshimoto A, Kishida Y, Arakawa R, Enomoto K (2008) Quorum sensing signaling molecules involved in the production of violacein by *Pseudoalteromonas*. Biosci Biotechnol Biochem. 72:1958-1961
50. Waters CM, Bassler BL (2005) Quorum sensing: cell-to-cell communication in bacteria. Annu Rev Cell Dev Biol. 21:319-346
51. Watson WT, Minogue TD, von Bodman SB, Val DL, Churchill MEA (2002) Structural basis and specificity of acyl-homoserine lactone signal production in bacterial quorum sensing. Mol Cell. 9:685-694

52. Whitehead NA, Barnard AM, Slater H, Simpson NJ, Salmond GP (2001) Quorum-sensing in Gram-negative bacteria. FEMS Microbiol Rev. 25:365-404
53. Williams P (2007) Quorum sensing, communication and cross-kingdom signalling in the bacterial world. Microbiology. 153:3923–3938
54. Yada S, Wang Y, Zou Y, Nagasaki K, Hosokawa K, Osaka I, Arakawa R, Enomoto K (2008) Isolation and characterization of two groups of novel marine bacteria producing violacein. Mar Biotechnol. 10:128-132
55. Yang LH, Xiong H, Lee OO, Qi SH, Qian PY (2007) Effect of agitation on violacein production in Pseudoalteromonas luteoviolacea isolated from a marine sponge. Lett Appl Microbiol. 44:625-30
56. Zhang X (2010) Violacein biosynthesis and its regulation in *Pseudoalteromonas* sp. 520P1. PhD Thesis. Kochi University of Technology
57. Zhang X, Enomoto K (2011) Characterization of a gene cluster and its putative promoter region for violacein biosynthesis in *Pseudoalteromonas* sp. 520P1. Appl Microbiol Biotechnol. 90:1963-1971

ABBREVIATIONS

Name	***Abbr.***
N-acylhomoserine lactone	AHL
N-(3-oxooctanoyl)-homoserine lactone	3-oxo-C8-HSL
N-tetradecanoyl-homoserine lactone	C14-HSL
5-bromo-4-chloro-3-indolyl-D-galactopyranoside	X-Gal
Isopropyl-β-D-thiogalactopyranoside	IPTG
TLC	Thin Layer Chromatography
Ampicillin	Amp
Kanamycin	Km
Gentamicin	Gm

APPENDIXES

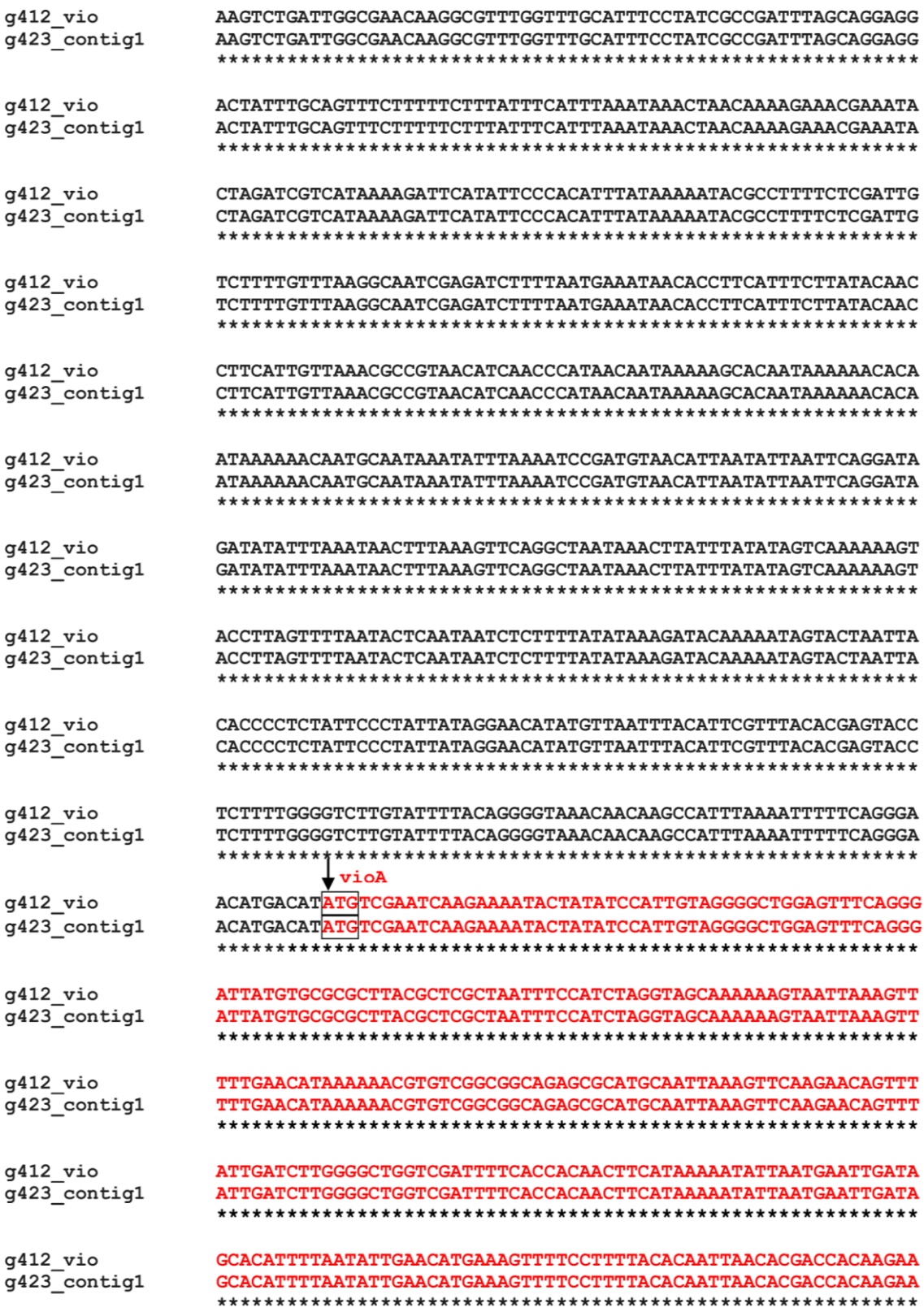
Appendix 1: IUPAC nucleotide code

IUPAC nucleotide code	Base
A	Adenine
C	Cytosine
G	Guanine
T (or U)	Thymine (or Uracil)
R	A or G
Y	C or T
S	G or C
W	A or T
K	G or T
M	A or C
B	C or G or T
D	A or G or T
H	A or C or T
V	A or C or G
N	any base

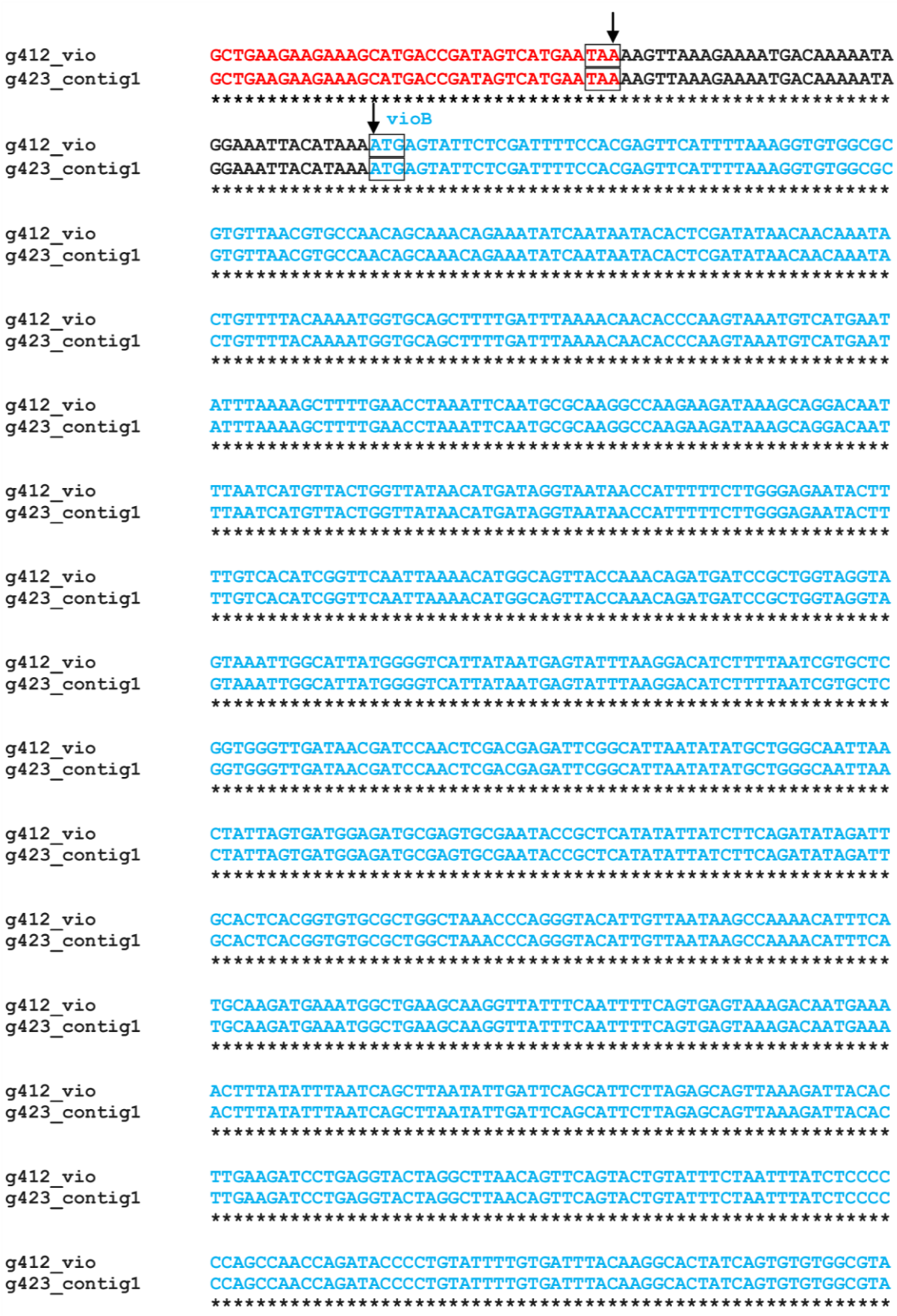
Appendix 2: Partial genomic analysis of strain 520P1 No. 412 & No. 423

* Violacein gene cluster

```
g412_vio          -GAAGCTAATAGCTACATTACCCCTATCATTATATTTTGGGTGATTGGTTTCATTAGCAA
g423_contig1      TGAAGCTAATAGCTACATTACCCCTATCATTATATTTTGGGTGATTGGTTTCATTAGCAA
                   ************************************************************

g412_vio          CCAAATAACCGAGGGATTTCGTGCTTCAGGAAATCCAAAAGCGATGATGCAGGTACTATC
g423_contig1      CCAAATAACCGAGGGATTTCGTGCTTCAGGAAATCCAAAAGCGATGATGCAGGTACTATC
                  ************************************************************

g412_vio          AACAGCTTCAATTCTCAACATCATTTTAGATGCGTTGTTTATATTTGTATTTGAATGGGG
g423_contig1      AACAGCTTCAATTCTCAACATCATTTTAGATGCGTTGTTTATATTTGTATTTGAATGGGG
                  ************************************************************

g412_vio          CATAGCTGGTGCTGCTTGGGCAACGATAATTGCAATTACACTAGGTTTATTAATGGCTAT
g423_contig1      CATAGCTGGTGCTGCTTGGGCAACGATAATTGCAATTACACTAGGTTTATTAATGGCTAT
                  ************************************************************

g412_vio          GCAATTACAGAAAAAAGGGGAAAGCGCGGTAAGATTTAGCTGGTGTAAGCTTATTAGCCC
g423_contig1      GCAATTACAGAAAAAAGGGGAAAGCGCGGTAAGATTTAGCTGGTGTAAGCTTATTAGCCC
                  ************************************************************

g412_vio          AATAAAGTTTCACCTAAAAATATTAGGTTTAGGATTGCCTGTATTATTATCCCACGGTGG
g423_contig1      AATAAAGTTTCACCTAAAAATATTAGGTTTAGGATTGCCTGTATTATTATCCCACGGTGG
                  ************************************************************

g412_vio          TTTTTCAGTGACGCTTGCAGTAACCGTTTATTCTATTTCAACTGTATTTATAGGTGTATC
g423_contig1      TTTTTCAGTGACGCTTGCAGTAACCGTTTATTCTATTTCAACTGTATTTATAGGTGTATC
                  ************************************************************

g412_vio          AGAGCCGTTAATCGCTGCACATGGCATATTAATAAGGTGCTTTATGTTCTTGTTTTTACC
g423_contig1      AGAGCCGTTAATCGCTGCACATGGCATATTAATAAGGTGCTTTATGTTCTTGTTTTTACC
                  ************************************************************

g412_vio          AATAATAGGCATGATGGTTGCCCTGCAAACCTTATCTGGATATAACTATGGCGCAGGTAA
g423_contig1      AATAATAGGCATGATGGTTGCCCTGCAAACCTTATCTGGATATAACTATGGCGCAGGTAA
                  ************************************************************

g412_vio          ATACCATAGGGTAAAACAGGCTTACTTTGTCGCTATTGCAACGAGTACTATTTGGGGCGC
g423_contig1      ATACCATAGGGTAAAACAGGCTTACTTTGTCGCTATTGCAACGAGTACTATTTGGGGCGC
                  ************************************************************

g412_vio          AATTGTAACATTCATTTTATGCTTCAATTCAGACTGGTTACTTACCATGTTTACAGATGA
g423_contig1      AATTGTAACATTCATTTTATGCTTCAATTCAGACTGGTTACTTACCATGTTTACAGATGA
                  ************************************************************

g412_vio          TATAGAGGTTATTGAACTTGGTAGTGAGCTAGCACCAATTTGTTTTGCTGGATTCATTAC
g423_contig1      TATAGAGGTTATTGAACTTGGTAGTGAGCTAGCACCAATTTGTTTTGCTGGATTCATTAC
                  ************************************************************

g412_vio          AGCGAGTTTTTGTATGATGTCTAGTGGACTATTTCAAGGTTTAGGTCGAGCTTTACCGGC
g423_contig1      AGCGAGTTTTTGTATGATGTCTAGTGGACTATTTCAAGGTTTAGGTCGAGCTTTACCGGC
                  ************************************************************

g412_vio          AACATTATTAGATGCAGCTAGAACCTATGTATTACTTTTACCATTGATGTACTTTTTACC
g423_contig1      AACATTATTAGATGCAGCTAGAACCTATGTATTACTTTTACCATTGATGTACTTTTTACC
                  ************************************************************
```

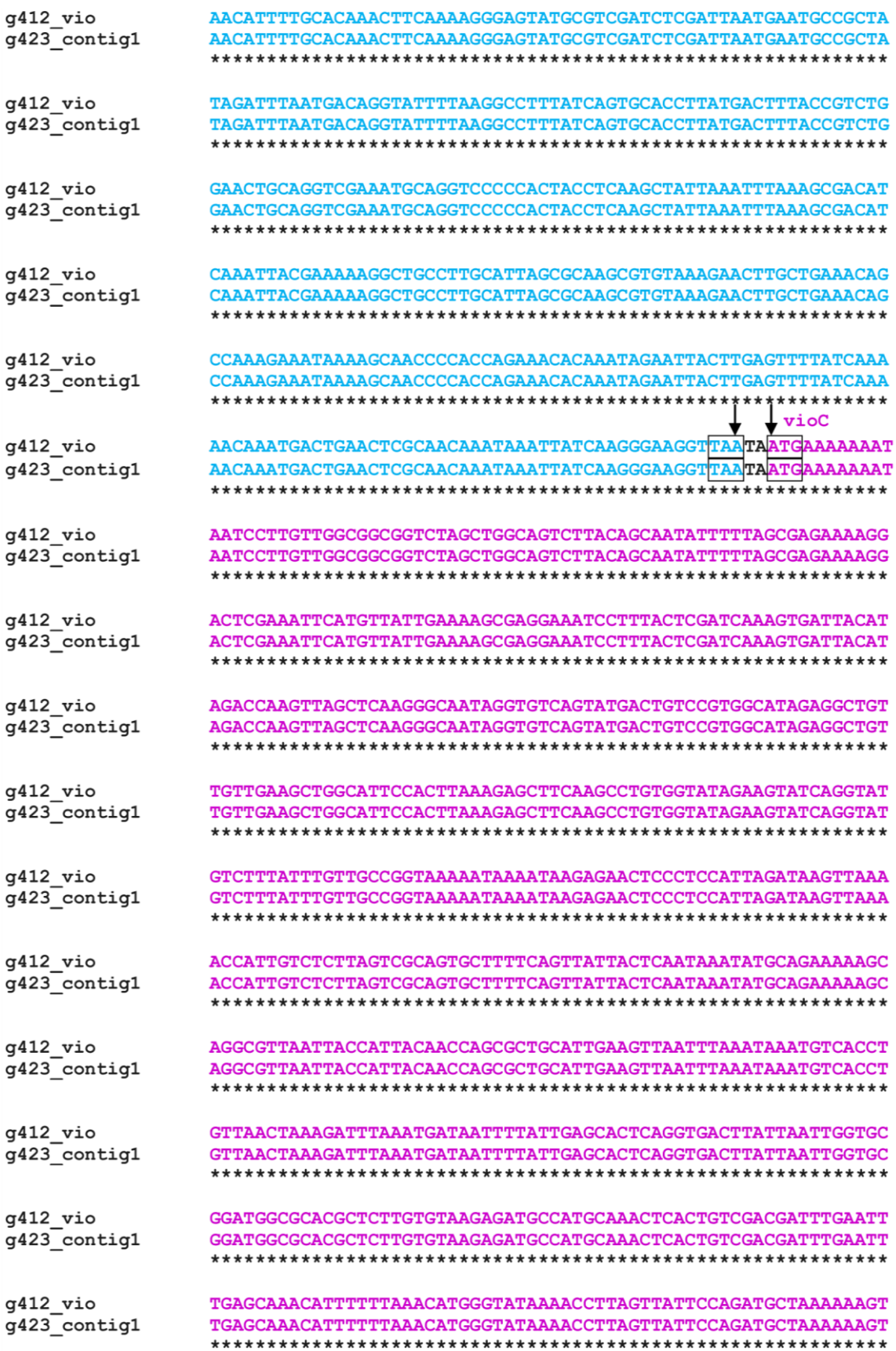

g412_vio AAGTCTGATTGGCGAACAAGGCGTTTGGTTTGCATTTCCTATCGCCGATTTAGCAGGAGG
g423_contig1 AAGTCTGATTGGCGAACAAGGCGTTTGGTTTGCATTTCCTATCGCCGATTTAGCAGGAGG
g412_vio ACTATTTGCAGTTTCTTTTTCTTTATTTCATTTAAATAAACTAACAAAAGAAACGAAATA
g423_contig1 ACTATTTGCAGTTTCTTTTTCTTTATTTCATTTAAATAAACTAACAAAAGAAACGAAATA
g412_vio CTAGATCGTCATAAAAGATTCATATTCCCACATTTATAAAAATACGCCTTTTCTCGATTG
g423_contig1 CTAGATCGTCATAAAAGATTCATATTCCCACATTTATAAAAATACGCCTTTTCTCGATTG
g412_vio TCTTTTGTTTAAGGCAATCGAGATCTTTTAATGAAATAACACCTTCATTTCTTATACAAC
g423_contig1 TCTTTTGTTTAAGGCAATCGAGATCTTTTAATGAAATAACACCTTCATTTCTTATACAAC
g412_vio CTTCATTGTTAAACGCCGTAACATCAACCCATAACAATAAAAAGCACAATAAAAAACACA
g423_contig1 CTTCATTGTTAAACGCCGTAACATCAACCCATAACAATAAAAAGCACAATAAAAAACACA
g412_vio ATAAAAAACAATGCAATAAATATTTAAAATCCGATGTAACATTAATATTAATTCAGGATA
g423_contig1 ATAAAAAACAATGCAATAAATATTTAAAATCCGATGTAACATTAATATTAATTCAGGATA
g412_vio GATATATTTAAATAACTTTAAAGTTCAGGCTAATAAACTTATTTATATAGTCAAAAAAGT
g423_contig1 GATATATTTAAATAACTTTAAAGTTCAGGCTAATAAACTTATTTATATAGTCAAAAAAGT
g412_vio ACCTTAGTTTTAATACTCAATAATCTCTTTTATATAAAGATACAAAAATAGTACTAATTA
g423_contig1 ACCTTAGTTTTAATACTCAATAATCTCTTTTATATAAAGATACAAAAATAGTACTAATTA
g412_vio CACCCCTCTATTCCCTATTATAGGAACATATGTTAATTTACATTCGTTTACACGAGTACC
g423_contig1 CACCCCTCTATTCCCTATTATAGGAACATATGTTAATTTACATTCGTTTACACGAGTACC
g412_vio TCTTTTGGGGTCTTGTATTTTACAGGGGTAAACAACAAGCCATTTAAAATTTTTCAGGGA
g423_contig1 TCTTTTGGGGTCTTGTATTTTACAGGGGTAAACAACAAGCCATTTAAAATTTTTCAGGGA
vioA
g412_vio ACATGACATATGTCGAATCAAGAAAATACTATATCCATTGTAGGGGCTGGAGTTTCAGGG
g423_contig1 ACATGACATATGTCGAATCAAGAAAATACTATATCCATTGTAGGGGCTGGAGTTTCAGGG
g412_vio ATTATGTGCGCGCTTACGCTCGCTAATTTCCATCTAGGTAGCAAAAAAGTAATTAAAGTT
g423_contig1 ATTATGTGCGCGCTTACGCTCGCTAATTTCCATCTAGGTAGCAAAAAAGTAATTAAAGTT
g412_vio TTTGAACATAAAAAACGTGTCGGCGGCAGAGCGCATGCAATTAAAGTTCAAGAACAGTTT
g423_contig1 TTTGAACATAAAAAACGTGTCGGCGGCAGAGCGCATGCAATTAAAGTTCAAGAACAGTTT
g412_vio ATTGATCTTGGGGCTGGTCGATTTTCACCACAACTTCATAAAAATATTAATGAATTGATA
g423_contig1 ATTGATCTTGGGGCTGGTCGATTTTCACCACAACTTCATAAAAATATTAATGAATTGATA
g412_vio GCACATTTTAATATTGAACATGAAAGTTTTCCTTTTACACAATTAACACGACCACAAGAA
g423_contig1 GCACATTTTAATATTGAACATGAAAGTTTTCCTTTTACACAATTAACACGACCACAAGAA

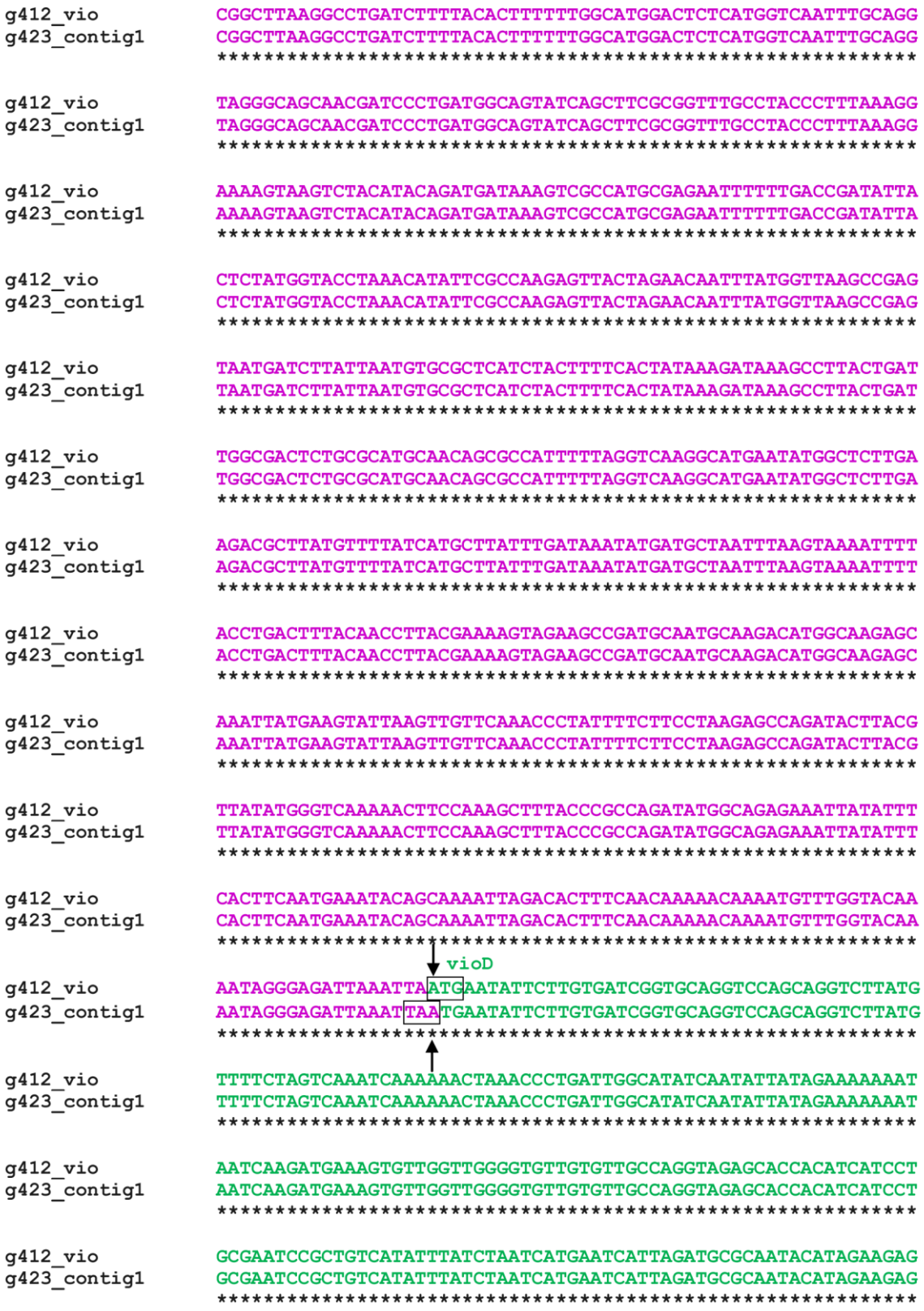
```
g412_vio          CTTCATAGTGAATTAAAGGAGATTTTAAACAAATTAAAGCCTTTGTGTGCTGAACACAAA
g423_contig1      CTTCATAGTGAATTAAAGGAGATTTTAAACAAATTAAAGCCTTTGTGTGCTGAACACAAA
                  ************************************************************

g412_vio          GATGAATCCTTTTTGGATTTTTTAACCTCTTATTTAGGTGAAGAACAAAGTAAAGATATT
g423_contig1      GATGAATCCTTTTTGGATTTTTTAACCTCTTATTTAGGTGAAGAACAAAGTAAAGATATT
                  ************************************************************

g412_vio          ATCAATGCGCTTGGATATGATTCTTTATCCTTGCCATTTATCACCCCCAATATTGCTTAC
g423_contig1      ATCAATGCGCTTGGATATGATTCTTTATCCTTGCCATTTATCACCCCCAATATTGCTTAC
                  ************************************************************

g412_vio          GATATTATTGAAAAGCACCCTGAGATCCAAGGGTTCTCTGATAACGCGGGATATACATGG
g423_contig1      GATATTATTGAAAAGCACCCTGAGATCCAAGGGTTCTCTGATAACGCGGGATATACATGG
                  ************************************************************

g412_vio          CGAAACCTTAAGGCCGGATTTTGCACTTTACCTCAAATCTTATATAAAAAAGCAGTTGAA
g423_contig1      CGAAACCTTAAGGCCGGATTTTGCACTTTACCTCAAATCTTATATAAAAAAGCAGTTGAA
                  ************************************************************

g412_vio          TTAGGTGTTGAGTTTCATTTTGATTATGAGCTTATTACAATTGATACCCATATAAAAAAA
g423_contig1      TTAGGTGTTGAGTTTCATTTTGATTATGAGCTTATTACAATTGATACCCATATAAAAAAA
                  ************************************************************

g412_vio          CCGGCTTTAACTTTAAAAAATAGTGATGATGAGATTATTCGAATAACCCAAGGGGATATG
g423_contig1      CCGGCTTTAACTTTAAAAAATAGTGATGATGAGATTATTCGAATAACCCAAGGGGATATG
                  ************************************************************

g412_vio          ATCGTCACATTGCCACCAACAGCAATGAGCAGGCTCAATTGTAACTTTCCGCAAGATTGG
g423_contig1      ATCGTCACATTGCCACCAACAGCAATGAGCAGGCTCAATTGTAACTTTCCGCAAGATTGG
                  ************************************************************

g412_vio          ACCCATTATACTTATGGTTCTATTCCTCTCTTCAAAGGTTTTCTTTTCTACGATACATCA
g423_contig1      ACCCATTATACTTATGGTTCTATTCCTCTCTTCAAAGGTTTTCTTTTCTACGATACATCA
                  ************************************************************

g412_vio          TGGTGGCAAGAATTATCATTAACAGATCATGTCATCATCACCAATAACCCAATACGAAAG
g423_contig1      TGGTGGCAAGAATTATCATTAACAGATCATGTCATCATCACCAATAACCCAATACGAAAG
                  ************************************************************

g412_vio          CTGTACTTTAAAGATAATAGATACATCTTTTTTTACACTGATAGTGAATACGCTAATTTC
g423_contig1      CTGTACTTTAAAGATAATAGATACATCTTTTTTTACACTGATAGTGAATACGCTAATTTC
                  ************************************************************

g412_vio          TGGCAAGAAACCAGTCAACATGGCGAAGAACATTACTTAAAAACAGTAAAAGAATTGATG
g423_contig1      TGGCAAGAAACCAGTCAACATGGCGAAGAACATTACTTAAAAACAGTAAAAGAATTGATG
                  ************************************************************

g412_vio          GCTGAAGCAATTCAGTGCCACATAGATCAAATTCCAGATCCAATTGAACAGACTCATAAA
g423_contig1      GCTGAAGCAATTCAGTGCCACATAGATCAAATTCCAGATCCAATTGAACAGACTCATAAA
                  ************************************************************

g412_vio          TATTGGATACATGGTGTTGAATTTTCCAAAGAGTGTAGTCCAACCCACCCTCTGAGTTTT
g423_contig1      TATTGGATACATGGTGTTGAATTTTCCAAAGAGTGTAGTCCAACCCACCCTCTGAGTTTT
                  ************************************************************

g412_vio          GTACACAAAAAAAGTAAAATAATTTCCGCTTCTGATGCTTATACACCCCACTGTGGTTGG
g423_contig1      GTACACAAAAAAAGTAAAATAATTTCCGCTTCTGATGCTTATACACCCCACTGTGGTTGG
                  ************************************************************

g412_vio          ATGGAAGGAGGCATTATGGCTGGTAAAAGTGCGGCCTTAAAACTTCTAAAACGGATTGAA
g423_contig1      ATGGAAGGAGGCATTATGGCTGGTAAAAGTGCGGCCTTAAAACTTCTAAAACGGATTGAA
                  ************************************************************
```


g412_vio
g423_contig1
GCTGAAGAAGAAAGCATGACCGATAGTCATGAATAAAAGTTAAAGAAAATGACAAAAATA
GCTGAAGAAGAAAGCATGACCGATAGTCATGAATAAAAGTTAAAGAAAATGACAAAAATA
vioB
GGAAATTACATAAAATGAGTATTCTCGATTTTCCACGAGTTCATTTTAAAGGTGTGGCGC
GGAAATTACATAAAATGAGTATTCTCGATTTTCCACGAGTTCATTTTAAAGGTGTGGCGC
GTGTTAACGTGCCAACAGCAAACAGAAATATCAATAATACACTCGATATAACAACAAATA
GTGTTAACGTGCCAACAGCAAACAGAAATATCAATAATACACTCGATATAACAACAAATA
CTGTTTTACAAAATGGTGCAGCTTTTGATTTAAAACAACACCCAAGTAAATGTCATGAAT
CTGTTTTACAAAATGGTGCAGCTTTTGATTTAAAACAACACCCAAGTAAATGTCATGAAT
ATTTAAAAGCTTTTGAACCTAAATTCAATGCGCAAGGCCAAGAAGATAAAGCAGGACAAT
ATTTAAAAGCTTTTGAACCTAAATTCAATGCGCAAGGCCAAGAAGATAAAGCAGGACAAT
TTAATCATGTTACTGGTTATAACATGATAGGTAATAACCATTTTTCTTGGGAGAATACTT
TTAATCATGTTACTGGTTATAACATGATAGGTAATAACCATTTTTCTTGGGAGAATACTT
TTGTCACATCGGTTCAATTAAAACATGGCAGTTACCAAACAGATGATCCGCTGGTAGGTA
TTGTCACATCGGTTCAATTAAAACATGGCAGTTACCAAACAGATGATCCGCTGGTAGGTA
GTAAATTGGCATTATGGGGTCATTATAATGAGTATTTAAGGACATCTTTTAATCGTGCTC
GTAAATTGGCATTATGGGGTCATTATAATGAGTATTTAAGGACATCTTTTAATCGTGCTC
GGTGGGTTGATAACGATCCAACTCGACGAGATTCGGCATTAATATATGCTGGGCAATTAA
GGTGGGTTGATAACGATCCAACTCGACGAGATTCGGCATTAATATATGCTGGGCAATTAA
CTATTAGTGATGGAGATGCGAGTGCGAATACCGCTCATATATTATCTTCAGATATAGATT
CTATTAGTGATGGAGATGCGAGTGCGAATACCGCTCATATATTATCTTCAGATATAGATT
GCACTCACGGTGTGCGCTGGCTAAACCCAGGGTACATTGTTAATAAGCCAAAACATTTCA
GCACTCACGGTGTGCGCTGGCTAAACCCAGGGTACATTGTTAATAAGCCAAAACATTTCA
TGCAAGATGAAATGGCTGAAGCAAGGTTATTTCAATTTTCAGTGAGTAAAGACAATGAAA
TGCAAGATGAAATGGCTGAAGCAAGGTTATTTCAATTTTCAGTGAGTAAAGACAATGAAA
ACTTTATATTTAATCAGCTTAATATTGATTCAGCATTCTTAGAGCAGTTAAAGATTACAC
ACTTTATATTTAATCAGCTTAATATTGATTCAGCATTCTTAGAGCAGTTAAAGATTACAC
TTGAAGATCCTGAGGTACTAGGCTTAACAGTTCAGTACTGTATTTCTAATTTATCTCCCC
TTGAAGATCCTGAGGTACTAGGCTTAACAGTTCAGTACTGTATTTCTAATTTATCTCCCC
CCAGCCAACCAGATACCCCTGTATTTTGTGATTTACAAGGCACTATCAGTGTGTGGCGTA
CCAGCCAACCAGATACCCCTGTATTTTGTGATTTACAAGGCACTATCAGTGTGTGGCGTA

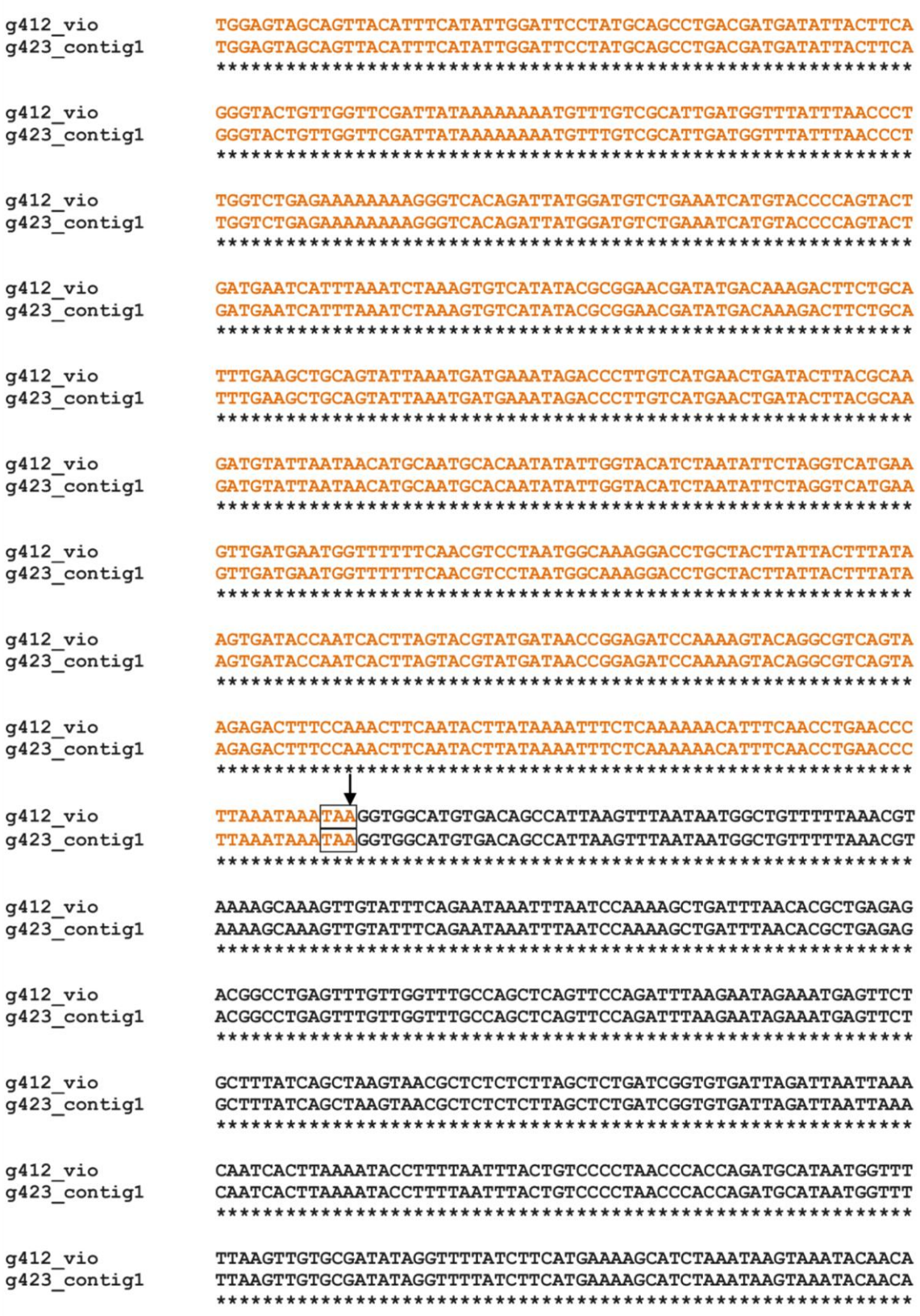
```
g412_vio          AACAAGATATGGCAACCAGCCCTACTGGCAGGATCCTACAACCAGATGATACCTCTCAAT
g423_contig1      AACAAGATATGGCAACCAGCCCTACTGGCAGGATCCTACAACCAGATGATACCTCTCAAT
                  ************************************************************

g412_vio          TTTCACCAATAGCGGTAAAAATAAAAGATAATTGGGTCAGTTTTAATATGCCCATTAGCA
g423_contig1      TTTCACCAATAGCGGTAAAAATAAAAGATAATTGGGTCAGTTTTAATATGCCCATTAGCA
                  ************************************************************

g412_vio          TTCCATATCAAAGTTATGCTGAGGTGTTGCCAGTTCAAAGCGGTCTGCCGCCAAAACTAA
g423_contig1      TTCCATATCAAAGTTATGCTGAGGTGTTGCCAGTTCAAAGCGGTCTGCCGCCAAAACTAA
                  ************************************************************

g412_vio          CCCACAAAGCCGCTCTTGGTGATTTAATATTAAAATCAGATTCAGGAAAAACGCTCGCTG
g423_contig1      CCCACAAAGCCGCTCTTGGTGATTTAATATTAAAATCAGATTCAGGAAAAACGCTCGCTG
                  ************************************************************

g412_vio          TTTTACCTGAATCCGTATACCAACAAGCCAATAACAGTACTGGTGTATTTGATGTGCCTT
g423_contig1      TTTTACCTGAATCCGTATACCAACAAGCCAATAACAGTACTGGTGTATTTGATGTGCCTT
                  ************************************************************

g412_vio          TACTCGTTAATGATGCGCAACTAGAAACGCAATCACTCAGTTTACAGTCAAATCAGCATA
g423_contig1      TACTCGTTAATGATGCGCAACTAGAAACGCAATCACTCAGTTTACAGTCAAATCAGCATA
                  ************************************************************

g412_vio          ACTGGCATGAAACTGATTGGCACATACAAGCTGAACAGCACATCATCGCAATTGAATCGG
g423_contig1      ACTGGCATGAAACTGATTGGCACATACAAGCTGAACAGCACATCATCGCAATTGAATCGG
                  ************************************************************

g412_vio          CAAATCCTAAATCTGATTATAAAAGTACACACGCTATCGATGTATTTAGTTATTTCAGAG
g423_contig1      CAAATCCTAAATCTGATTATAAAAGTACACACGCTATCGATGTATTTAGTTATTTCAGAG
                  ************************************************************

g412_vio          GTAAACCCCACCCTATTAATAAATTAATTCCTAATATAACAACGCCTAACAATTTACGGT
g423_contig1      GTAAACCCCACCCTATTAATAAATTAATTCCTAATATAACAACGCCTAACAATTTACGGT
                  ************************************************************

g412_vio          GTGATGTTTATATCGAAACTGATAAAAGTGGTAGAGGTCAACTCAATATTGAAAGCTTAG
g423_contig1      GTGATGTTTATATCGAAACTGATAAAAGTGGTAGAGGTCAACTCAATATTGAAAGCTTAG
                  ************************************************************

g412_vio          CGCCTGGTAGCGGTGAATTGTTTTTAGGTGAGCACCATAGCCCAGTACAAGTTCGCATAC
g423_contig1      CGCCTGGTAGCGGTGAATTGTTTTTAGGTGAGCACCATAGCCCAGTACAAGTTCGCATAC
                  ************************************************************

g412_vio          TCAGTGATGATTGGTCTTTACTCGATGTCGCAGATGAAAATGTAGATTATGATTTCCTTT
g423_contig1      TCAGTGATGATTGGTCTTTACTCGATGTCGCAGATGAAAATGTAGATTATGATTTCCTTT
                  ************************************************************

g412_vio          ATCATAATGTAATGGGCTATTACGAACTACTTTATCCTTTTATGGCTGACAAAGTATTTA
g423_contig1      ATCATAATGTAATGGGCTATTACGAACTACTTTATCCTTTTATGGCTGACAAAGTATTTA
                  ************************************************************

g412_vio          GTATGGCCGATAAGTGTAAATGTGAGACTTATGCCCGTTTAATGTGGCAAATGTGCGACC
g423_contig1      GTATGGCCGATAAGTGTAAATGTGAGACTTATGCCCGTTTAATGTGGCAAATGTGCGACC
                  ************************************************************

g412_vio          CTAAAAACCGCGACAAAAGTTATTACATGCCTAGTACTCGAGAAATGTCCTCGGTTAAGT
g423_contig1      CTAAAAACCGCGACAAAAGTTATTACATGCCTAGTACTCGAGAAATGTCCTCGGTTAAGT
                  ************************************************************

g412_vio          CACATTTATTTCTGAAATATTTAAGTAATGTTGAGCAATCAGCAATGCCTAAAGAGCTAC
g423_contig1      CACATTTATTTCTGAAATATTTAAGTAATGTTGAGCAATCAGCAATGCCTAAAGAGCTAC
                  ************************************************************
```

```
g412_vio          CGCCATTAGAGCCCCAGTTTACTGCGCAAGGTAGTATTAAAACTAAAGCTCAGTTAATAA
g423_contig1      CGCCATTAGAGCCCCAGTTTACTGCGCAAGGTAGTATTAAAACTAAAGCTCAGTTAATAA
                  ************************************************************

g412_vio          GCAAGCTACGTGATGCGGTAGATTTAGAGTTATCGATTATGTTGCAATATCTTTATAGCG
g423_contig1      GCAAGCTACGTGATGCGGTAGATTTAGAGTTATCGATTATGTTGCAATATCTTTATAGCG
                  ************************************************************

g412_vio          CCTATTCGTTACCTACCTATGCTGCAGGGGAGCAATATGTAGAGTCAGAACGTTGGACAC
g423_contig1      CCTATTCGTTACCTACCTATGCTGCAGGGGAGCAATATGTAGAGTCAGAACGTTGGACAC
                  ************************************************************

g412_vio          AAGCTCAGTTAGAGTTAGTTAACGGTTCCAAGGAAAGGCGAAAAAACAGTGGTTGGCGAG
g423_contig1      AAGCTCAGTTAGAGTTAGTTAACGGTTCCAAGGAAAGGCGAAAAAACAGTGGTTGGCGAG
                  ************************************************************

g412_vio          GTGCTATTTTAGAGATCGCCCATGAAGAAATGATACATTATTTGGTCATCAATAATATCT
g423_contig1      GTGCTATTTTAGAGATCGCCCATGAAGAAATGATACATTATTTGGTCATCAATAATATCT
                  ************************************************************

g412_vio          TGATGTCTCTAGGTGAGCCTTTCTACCCTGGAGAGCCTGTTTTTGCACAAGCCGCCAAAG
g423_contig1      TGATGTCTCTAGGTGAGCCTTTCTACCCTGGAGAGCCTGTTTTTGCACAAGCCGCCAAAG
                  ************************************************************

g412_vio          AGAAGTTTGGCTTAGATACTGAATTTTCTTTTGAGCCTTTTTCTGAGCATATTATTGCTA
g423_contig1      AGAAGTTTGGCTTAGATACTGAATTTTCTTTTGAGCCTTTTTCTGAGCATATTATTGCTA
                  ************************************************************

g412_vio          AGTTTGTCCGCTTTGAATGGCCTCATTTCTTTCCTTCTGTTGGTAAGTCGATAGCCGATT
g423_contig1      AGTTTGTCCGCTTTGAATGGCCTCATTTCTTTCCTTCTGTTGGTAAGTCGATAGCCGATT
                  ************************************************************

g412_vio          TTTATAACGAAATTCGTATTGCTATCAACGAAATACCCGACTTATATACCCAAGACATGA
g423_contig1      TTTATAACGAAATTCGTATTGCTATCAACGAAATACCCGACTTATATACCCAAGACATGA
                  ************************************************************

g412_vio          ATAAGCAAGGTGGTGAACATCACTTATTTTTAAACGAAATAATAAACCGTGCTTATCCTA
g423_contig1      ATAAGCAAGGTGGTGAACATCACTTATTTTTAAACGAAATAATAAACCGTGCTTATCCTA
                  ************************************************************

g412_vio          ATTATCAGTTTGAAGTTTATGATAAAGAAACTGCATTATTTGCTATCGACTTTGTTACCG
g423_contig1      ATTATCAGTTTGAAGTTTATGATAAAGAAACTGCATTATTTGCTATCGACTTTGTTACCG
                  ************************************************************

g412_vio          AGCAAGGTGAAGGCGCTAGTGCTGATTCTCCACAATTTGAACATAGCCATTTTAACCGCT
g423_contig1      AGCAAGGTGAAGGCGCTAGTGCTGATTCTCCACAATTTGAACATAGCCATTTTAACCGCT
                  ************************************************************

g412_vio          TAAGATCTATATCTAAAAACCTCACTCTTAGCGACATTCCTTTTGAACCTGCTTATCCCG
g423_contig1      TAAGATCTATATCTAAAAACCTCACTCTTAGCGACATTCCTTTTGAACCTGCTTATCCCG
                  ************************************************************

g412_vio          TTTTAAAAAATCCAGTGATAAGTCAGCGAGCAGGGTGCAATGTTGTTACAAACCCGAATG
g423_contig1      TTTTAAAAAATCCAGTGATAAGTCAGCGAGCAGGGTGCAATGTTGTTACAAACCCGAATG
                  ************************************************************

g412_vio          CCAGAGCTTTAATGACGCTTTATCAAGGTTGTCATGAACTGATGTTTAAAATGATGATGC
g423_contig1      CCAGAGCTTTAATGACGCTTTATCAAGGTTGTCATGAACTGATGTTTAAAATGATGATGC
                  ************************************************************
```


g412_vio AACATTTTGCACAAACTTCAAAAGGGAGTATGCGTCGATCTCGATTAATGAATGCCGCTA
g423_contig1 AACATTTTGCACAAACTTCAAAAGGGAGTATGCGTCGATCTCGATTAATGAATGCCGCTA
g412_vio TAGATTTAATGACAGGTATTTTAAGGCCTTTATCAGTGCACCTTATGACTTTACCGTCTG
g423_contig1 TAGATTTAATGACAGGTATTTTAAGGCCTTTATCAGTGCACCTTATGACTTTACCGTCTG
g412_vio GAACTGCAGGTCGAAATGCAGGTCCCCCACTACCTCAAGCTATTAAATTTAAAGCGACAT
g423_contig1 GAACTGCAGGTCGAAATGCAGGTCCCCCACTACCTCAAGCTATTAAATTTAAAGCGACAT
g412_vio CAAATTACGAAAAAGGCTGCCTTGCATTAGCGCAAGCGTGTAAAGAACTTGCTGAAACAG
g423_contig1 CAAATTACGAAAAAGGCTGCCTTGCATTAGCGCAAGCGTGTAAAGAACTTGCTGAAACAG
g412_vio CCAAAGAAATAAAAGCAACCCCACCAGAAACACAAATAGAATTACTTGAGTTTTATCAAA
g423_contig1 CCAAAGAAATAAAAGCAACCCCACCAGAAACACAAATAGAATTACTTGAGTTTTATCAAA
vioC
g412_vio AACAAATGACTGAACTCGCAACAAATAAATTATCAAGGGAAGGTTAATAATGAAAAAAAT
g423_contig1 AACAAATGACTGAACTCGCAACAAATAAATTATCAAGGGAAGGTTAATAATGAAAAAAAT
g412_vio AATCCTTGTTGGCGGCGGTCTAGCTGGCAGTCTTACAGCAATATTTTTAGCGAGAAAAGG
g423_contig1 AATCCTTGTTGGCGGCGGTCTAGCTGGCAGTCTTACAGCAATATTTTTAGCGAGAAAAGG
g412_vio ACTCGAAATTCATGTTATTGAAAAGCGAGGAAATCCTTTACTCGATCAAAGTGATTACAT
g423_contig1 ACTCGAAATTCATGTTATTGAAAAGCGAGGAAATCCTTTACTCGATCAAAGTGATTACAT
g412_vio AGACCAAGTTAGCTCAAGGGCAATAGGTGTCAGTATGACTGTCCGTGGCATAGAGGCTGT
g423_contig1 AGACCAAGTTAGCTCAAGGGCAATAGGTGTCAGTATGACTGTCCGTGGCATAGAGGCTGT
g412_vio TGTTGAAGCTGGCATTCCACTTAAAGAGCTTCAAGCCTGTGGTATAGAAGTATCAGGTAT
g423_contig1 TGTTGAAGCTGGCATTCCACTTAAAGAGCTTCAAGCCTGTGGTATAGAAGTATCAGGTAT
g412_vio GTCTTTATTTGTTGCCGGTAAAAATAAAATAAGAGAACTCCCTCCATTAGATAAGTTAAA
g423_contig1 GTCTTTATTTGTTGCCGGTAAAAATAAAATAAGAGAACTCCCTCCATTAGATAAGTTAAA
g412_vio ACCATTGTCTCTTAGTCGCAGTGCTTTTCAGTTATTACTCAATAAATATGCAGAAAAAGC
g423_contig1 ACCATTGTCTCTTAGTCGCAGTGCTTTTCAGTTATTACTCAATAAATATGCAGAAAAAGC
g412_vio AGGCGTTAATTACCATTACAACCAGCGCTGCATTGAAGTTAATTTAAATAAATGTCACCT
g423_contig1 AGGCGTTAATTACCATTACAACCAGCGCTGCATTGAAGTTAATTTAAATAAATGTCACCT
g412_vio GTTAACTAAAGATTTAAATGATAATTTTATTGAGCACTCAGGTGACTTATTAATTGGTGC
g423_contig1 GTTAACTAAAGATTTAAATGATAATTTTATTGAGCACTCAGGTGACTTATTAATTGGTGC
g412_vio GGATGGCGCACGCTCTTGTGTAAGAGATGCCATGCAAACTCACTGTCGACGATTTGAATT
g423_contig1 GGATGGCGCACGCTCTTGTGTAAGAGATGCCATGCAAACTCACTGTCGACGATTTGAATT
g412_vio TGAGCAAACATTTTTTAAACATGGGTATAAAACCTTAGTTATTCCAGATGCTAAAAAAGT
g423_contig1 TGAGCAAACATTTTTTAAACATGGGTATAAAACCTTAGTTATTCCAGATGCTAAAAAAGT


g412_vio
g423_contig1
CGGCTTAAGGCCTGATCTTTTACACTTTTTTGGCATGGACTCTCATGGTCAATTTGCAGG
CGGCTTAAGGCCTGATCTTTTACACTTTTTTGGCATGGACTCTCATGGTCAATTTGCAGG
TAGGGCAGCAACGATCCCTGATGGCAGTATCAGCTTCGCGGTTTGCCTACCCTTTAAAGG
TAGGGCAGCAACGATCCCTGATGGCAGTATCAGCTTCGCGGTTTGCCTACCCTTTAAAGG
AAAAGTAAGTCTACATACAGATGATAAAGTCGCCATGCGAGAATTTTTTGACCGATATTA
AAAAGTAAGTCTACATACAGATGATAAAGTCGCCATGCGAGAATTTTTTGACCGATATTA
CTCTATGGTACCTAAACATATTCGCCAAGAGTTACTAGAACAATTTATGGTTAAGCCGAG
CTCTATGGTACCTAAACATATTCGCCAAGAGTTACTAGAACAATTTATGGTTAAGCCGAG
TAATGATCTTATTAATGTGCGCTCATCTACTTTTCACTATAAAGATAAAGCCTTACTGAT
TAATGATCTTATTAATGTGCGCTCATCTACTTTTCACTATAAAGATAAAGCCTTACTGAT
TGGCGACTCTGCGCATGCAACAGCGCCATTTTTAGGTCAAGGCATGAATATGGCTCTTGA
TGGCGACTCTGCGCATGCAACAGCGCCATTTTTAGGTCAAGGCATGAATATGGCTCTTGA
AGACGCTTATGTTTTATCATGCTTATTTGATAAATATGATGCTAATTTAAGTAAAATTTT
AGACGCTTATGTTTTATCATGCTTATTTGATAAATATGATGCTAATTTAAGTAAAATTTT
ACCTGACTTTACAACCTTACGAAAAGTAGAAGCCGATGCAATGCAAGACATGGCAAGAGC
ACCTGACTTTACAACCTTACGAAAAGTAGAAGCCGATGCAATGCAAGACATGGCAAGAGC
AAATTATGAAGTATTAAGTTGTTCAAACCCTATTTTCTTCCTAAGAGCCAGATACTTACG
AAATTATGAAGTATTAAGTTGTTCAAACCCTATTTTCTTCCTAAGAGCCAGATACTTACG
TTATATGGGTCAAAAACTTCCAAAGCTTTACCCGCCAGATATGGCAGAGAAATTATATTT
TTATATGGGTCAAAAACTTCCAAAGCTTTACCCGCCAGATATGGCAGAGAAATTATATTT
CACTTCAATGAAATACAGCAAAATTAGACACTTTCAACAAAAACAAAATGTTTGGTACAA
CACTTCAATGAAATACAGCAAAATTAGACACTTTCAACAAAAACAAAATGTTTGGTACAA
vioD
AATAGGGAGATTAAATTAATGAATATTCTTGTGATCGGTGCAGGTCCAGCAGGTCTTATG
AATAGGGAGATTAAATTAATGAATATTCTTGTGATCGGTGCAGGTCCAGCAGGTCTTATG
TTTTCTAGTCAAATCAAAAAACTAAACCCTGATTGGCATATCAATATTATAGAAAAAAAT
TTTTCTAGTCAAATCAAAAAACTAAACCCTGATTGGCATATCAATATTATAGAAAAAAAT
AATCAAGATGAAAGTGTTGGTTGGGGTGTTGTGTTGCCAGGTAGAGCACCACATCATCCT
AATCAAGATGAAAGTGTTGGTTGGGGTGTTGTGTTGCCAGGTAGAGCACCACATCATCCT
GCGAATCCGCTGTCATATTTATCTAATCATGAATCATTAGATGCGCAATACATAGAAGAG
GCGAATCCGCTGTCATATTTATCTAATCATGAATCATTAGATGCGCAATACATAGAAGAG

```
g412_vio          TTTAAGCTTACACATCATAATGACTCAGCCTTAACTAAAACCGGCGTCACTTTATGTGGT
g423_contig1      TTTAAGCTTACACATCATAATGACTCAGCCTTAACTAAAACCGGCGTCACTTTATGTGGT
                  ************************************************************

g412_vio          GCCGAGCGTAAATCTATGGTGCATGAACTTCGTCAATTATGTATTGGTTTAGGTATTTCG
g423_contig1      GCCGAGCGTAAATCTATGGTGCATGAACTTCGTCAATTATGTATTGGTTTAGGTATTTCG
                  ************************************************************

g412_vio          ATTGAATATGAAAAACCGGCATCAAAACTCGTTGACCTACAATGTAACAAATATGATTTA
g423_contig1      ATTGAATATGAAAAACCGGCATCAAAACTCGTTGACCTACAATGTAACAAATATGATTTA
                  ************************************************************

g412_vio          GTTGTTGTTTCAAATGGTATTAATCATACATCAACGTACTATAAAGAAGCATTAAAGCCT
g423_contig1      GTTGTTGTTTCAAATGGTATTAATCATACATCAACGTACTATAAAGAAGCATTAAAGCCT
                  ************************************************************

g412_vio          AAAGTTGAATTCGGTAAAAATCGTTATATGTGGTACGGGACGACTAAAAAGTTTGATGAA
g423_contig1      AAAGTTGAATTCGGTAAAAATCGTTATATGTGGTACGGGACGACTAAAAAGTTTGATGAA
                  ************************************************************

g412_vio          ATGAATTTAATTTTCAAAACGAAAGCTAAAGGTATTTTTGTTGCTCACTGTTATAAATAT
g423_contig1      ATGAATTTAATTTTCAAAACGAAAGCTAAAGGTATTTTTGTTGCTCACTGTTATAAATAT
                  ************************************************************

g412_vio          TCCAGTAATATGAGTACTTTTGTTGTTGAATGTAGTGAAGAAACGTATATAAATTCAGGT
g423_contig1      TCCAGTAATATGAGTACTTTTGTTGTTGAATGTAGTGAAGAAACGTATATAAATTCAGGT
                  ************************************************************

g412_vio          CTTGATGAAATGTCGACTCAAAATGCCGAAGCCTTTATTGCTTCAGTATTTGAAGAAGAG
g423_contig1      CTTGATGAAATGTCGACTCAAAATGCCGAAGCCTTTATTGCTTCAGTATTTGAAGAAGAG
                  ************************************************************

g412_vio          TTAGACGGTCAAACGGTTATATCACCAAAAGGGCTCAAATGGCGTAACTTCATGACATTG
g423_contig1      TTAGACGGTCAAACGGTTATATCACCAAAAGGGCTCAAATGGCGTAACTTCATGACATTG
                  ************************************************************

g412_vio          AGTCATGAACAAGCCTATAGCGATAATATTGTTTTACTAGGTGATGCACTGCAATCAGGA
g423_contig1      AGTCATGAACAAGCCTATAGCGATAATATTGTTTTACTAGGTGATGCACTGCAATCAGGA
                  ************************************************************

g412_vio          CATTTTTCTATTGGTCACGGTACGACTATGGCTGTAGTTGGCGCACAAATGTTAGTTAAA
g423_contig1      CATTTTTCTATTGGTCACGGTACGACTATGGCTGTAGTTGGCGCACAAATGTTAGTTAAA
                  ************************************************************

g412_vio          TCGGTTTACGATCATAGTGATAATATTGCTACGGCGTTAGAAGATTTTAACCAAAATGTG
g423_contig1      TCGGTTTACGATCATAGTGATAATATTGCTACGGCGTTAGAAGATTTTAACCAAAATGTG
                  ************************************************************

g412_vio          ATGCCTGTCATGCAATTGTTTGATCAACATGCAAGTACTAGCCGATTATGGTTTGAAAGC
g423_contig1      ATGCCTGTCATGCAATTGTTTGATCAACATGCAAGTACTAGCCGATTATGGTTTGAAAGC
                  ************************************************************

g412_vio          GTAGAAGACCGCATGCATTTATCAACTCCTGAGCTAGCACAAAGCTTTGCGACGCGCAGA
g423_contig1      GTAGAAGACCGCATGCATTTATCAACTCCTGAGCTAGCACAAAGCTTTGCGACGCGCAGA
                  ************************************************************

g412_vio          AACCAATTACCTCCTCTACCGCCAGCGTTAGGACAAGCACTTGAAAAAGCTTTAGCGCGA
g423_contig1      AACCAATTACCTCCTCTACCGCCAGCGTTAGGACAAGCACTTGAAAAAGCTTTAGCGCGA
                  ************************************************************
                            vioE
g412_vio          GGAGAAAAGTAAATGCAGTTTAATAAAATAACAAAAGTCGCTCCTTTATTACCTGAACAA
g423_contig1      GGAGAAAAGTAAATGCAGTTTAATAAAATAACAAAAGTCGCTCCTTTATTACCTGAACAA
                  ************************************************************
```


g412_vio TGGAGTAGCAGTTACATTTCATATTGGATTCCTATGCAGCCTGACGATGATATTACTTCA
g423_contig1 TGGAGTAGCAGTTACATTTCATATTGGATTCCTATGCAGCCTGACGATGATATTACTTCA
g412_vio GGGTACTGTTGGTTCGATTATAAAAAAAATGTTTGTCGCATTGATGGTTTATTTAACCCT
g423_contig1 GGGTACTGTTGGTTCGATTATAAAAAAAATGTTTGTCGCATTGATGGTTTATTTAACCCT
g412_vio TGGTCTGAGAAAAAAAAGGGTCACAGATTATGGATGTCTGAAATCATGTACCCCAGTACT
g423_contig1 TGGTCTGAGAAAAAAAAGGGTCACAGATTATGGATGTCTGAAATCATGTACCCCAGTACT
g412_vio GATGAATCATTTAAATCTAAAGTGTCATATACGCGGAACGATATGACAAAGACTTCTGCA
g423_contig1 GATGAATCATTTAAATCTAAAGTGTCATATACGCGGAACGATATGACAAAGACTTCTGCA
g412_vio TTTGAAGCTGCAGTATTAAATGATGAAATAGACCCTTGTCATGAACTGATACTTACGCAA
g423_contig1 TTTGAAGCTGCAGTATTAAATGATGAAATAGACCCTTGTCATGAACTGATACTTACGCAA
g412_vio GATGTATTAATAACATGCAATGCACAATATATTGGTACATCTAATATTCTAGGTCATGAA
g423_contig1 GATGTATTAATAACATGCAATGCACAATATATTGGTACATCTAATATTCTAGGTCATGAA
g412_vio GTTGATGAATGGTTTTTTCAACGTCCTAATGGCAAAGGACCTGCTACTTATTACTTTATA
g423_contig1 GTTGATGAATGGTTTTTTCAACGTCCTAATGGCAAAGGACCTGCTACTTATTACTTTATA
g412_vio AGTGATACCAATCACTTAGTACGTATGATAACCGGAGATCCAAAAGTACAGGCGTCAGTA
g423_contig1 AGTGATACCAATCACTTAGTACGTATGATAACCGGAGATCCAAAAGTACAGGCGTCAGTA
g412_vio AGAGACTTTCCAAACTTCAATACTTATAAAATTTCTCAAAAAACATTTCAACCTGAACCC
g423_contig1 AGAGACTTTCCAAACTTCAATACTTATAAAATTTCTCAAAAAACATTTCAACCTGAACCC
g412_vio TTAAATAAATAAGGTGGCATGTGACAGCCATTAAGTTTAATAATGGCTGTTTTTAAACGT
g423_contig1 TTAAATAAATAAGGTGGCATGTGACAGCCATTAAGTTTAATAATGGCTGTTTTTAAACGT
g412_vio AAAAGCAAAGTTGTATTTCAGAATAAATTTAATCCAAAAGCTGATTTAACACGCTGAGAG
g423_contig1 AAAAGCAAAGTTGTATTTCAGAATAAATTTAATCCAAAAGCTGATTTAACACGCTGAGAG
g412_vio ACGGCCTGAGTTTGTTGGTTTGCCAGCTCAGTTCCAGATTTAAGAATAGAAATGAGTTCT
g423_contig1 ACGGCCTGAGTTTGTTGGTTTGCCAGCTCAGTTCCAGATTTAAGAATAGAAATGAGTTCT
g412_vio GCTTTATCAGCTAAGTAACGCTCTCTCTTAGCTCTGATCGGTGTGATTAGATTAATTAAA
g423_contig1 GCTTTATCAGCTAAGTAACGCTCTCTCTTAGCTCTGATCGGTGTGATTAGATTAATTAAA
g412_vio CAATCACTTAAAATACCTTTTAATTTACTGTCCCCTAACCCACCAGATGCATAATGGTTT
g423_contig1 CAATCACTTAAAATACCTTTTAATTTACTGTCCCCTAACCCACCAGATGCATAATGGTTT
g412_vio TTAAGTTGTGCGATATAGGTTTTATCTTCATGAAAAGCATCTAAATAAGTAAATACAACA
g423_contig1 TTAAGTTGTGCGATATAGGTTTTATCTTCATGAAAAGCATCTAAATAAGTAAATACAACA

```
g412_vio          TTACCTTCTACTTGCCCAGGATCGTCTACTCTTAAGTGATTTGAATCGGTATACATAGCG
g423_contig1      TTACCTTCTACTTGCCCAGGATCGTCTACTCTTAAGTGATTTGAATCGGTATACATAGCG
                  ************************************************************

g412_vio          CGAACAGCTTTTACTATCATGTCATCACTACTACCAAAAGTAATCACATTACCCAGTGAT
g423_contig1      CGAACAGCTTTTACTATCATGTCATCACTACTACCAAAAGTAATCACATTACCCAGTGAT
                  ************************************************************


g412_vio          TTAGACATTTTAGCTTTACCATCTACACCCGGTAATCTAGGTGTATCACTCAATAATGGT
g423_contig1      TTAGACATTTTAGCTTTACCATCTACACCCGGTAATCTAGGTGTATCACTCAATAATGGT
                  ************************************************************

g412_vio          TTCGCTTCAATTAAAACAGGCTGCTGTGCAATATTGTTTAATTTTCTTACTATTTCATTG
g423_contig1      TTCGCTTCAATTAAAACAGGCTGCTGTGCAATATTGTTTAATTTTCTTACTATTTCATTG
                  ************************************************************

g412_vio          GTAATTTCTATCATAGGTAACTGATCATCGCCAACAGGAACTAAAGTCGCATCAAAAGCT
g423_contig1      GTAATTTCTATCATAGGTAACTGATCATCGCCAACAGGAACTAAAGTCGCATCAAAAGCT
                  ************************************************************

g412_vio          GTAATATCGGCAGCTTG----------------------------------------
g423_contig1      GTAATATCGGCAGCTTGGCTAATTGGGTAACTAAAAAAACCAGCAGGAATAGAGTTACCA
                  *****************
```

*** *palI* and its promoter**

```
g423_contig1          TTCCAAAAACTTGTTTTTTGAACACCTGATACATCTGACAGCAATTCCATTATTTAAACA
g412_scaffold19       TTCCAAAAACTTGTTTTTTGAACACCTGATACATCTGACAGCAATTCCATTATTTAAACA
palI                  ------------------------------------------------------------


g423_contig1          TTAATAAATTTCAGATACACATTAAAAACACTACTTAACAGATACTTAAGTTTAATTTTA
g412_scaffold19       TTAATAAATTTCAGATACACATTAAAAACACTACTTAACAGATACTTAAGTTTAATTTTA
palI                  ------------------------------------------------------------


g423_contig1          AGAACTTTAATCAAATAAGAATTTGGAAGGTTTTTTAAATATAGATAAAACAAAAAGATA
g412_scaffold19       AGAACTTTAATCAAATAAGAATTTGGAAGGTTTTTTAAATATAGATAAAACAAAAAGATA
palI                  ------------------------------------------------------------


g423_contig1          CCCTGTAACTTTTTACAGGTTTTGGGTACAAGATAGTTTGTTATAGTGATTTAAAAGAAA
g412_scaffold19       CCCTGTAACTTTTTACAGGTTTTGGGTACAAGATAGTTTGTTATAGTGATTTAAAAGAAA
palI                  ------------------------------------------------------------


g423_contig1          TGTAAATCAGAATAATTATTAATGAATCTCTATAACAATCAACTCTCATCTATTGATAAA
g412_scaffold19       TGTAAATCAGAATAATTATTAATGAATCTCTATAACAATCAACTCTCATCTATTGATAAA
palI                  --------------------ATGAATCTCTATAACAATCAACTCTCATCTATTGATAAA
                                          ************************************


g423_contig1          AGTAAAATTATTGCAATGCACCAATTAAGAGAAAAGGCTTTTTATAAACGCTTAAAATGG
g412_scaffold19       AGTAAAATTATTGCAATGCACCAATTAAGAGAAAAGGCTTTTTATAAACGCTTAAAATGG
palI                  AGTAAAATTATTGCAATGCACCAATTAAGAGAAAAGGCTTTTTATAAACGCTTAAAATGG
                      ************************************************************


g423_contig1          CAAGTAGATAGTTTTAATGGTATGGAAAAAGATAAATTTGATAATATTTACTCTTCATAT
g412_scaffold19       CAAGTAGATAGTTTTAATGGTATGGAAAAAGATAAATTTGATAATATTTACTCTTCATAT
palI                  CAAGTAGATAGTTTTAATGGTATGGAAAAAGATAAATTTGATAATATTTACTCTTCATAT
                      ************************************************************


g423_contig1          TTACTTGTCAAATCTCAATCAAATGAAGTAATAGGTTGTGCTAGATTAATTCCTACTGAT
g412_scaffold19       TTACTTGTCAAATCTCAATCAAATGAAGTAATAGGTTGTGCTAGATTAATTCCTACTGAT
palI                  TTACTTGTCAAATCTCAATCAAATGAAGTAATAGGTTGTGCTAGATTAATTCCTACTGAT
                      ************************************************************


g423_contig1          CAAGATTACATGTTTAAAAACATATTTTCAGAACTTGCTAGAAATGAAGAAATCCCAGAA
g412_scaffold19       CAAGATTACATGTTTAAAAACATATTTTCAGAACTTGCTAGAAATGAAGAAATCCCAGAA
palI                  CAAGATTACATGTTTAAAAACATATTTTCAGAACTTGCTAGAAATGAAGAAATCCCAGAA
                      ************************************************************


g423_contig1          AAATCAGATGTATGGGAAATAAGTCGATTTACATTAGATCGTACATTTGAACGAAACTAC
g412_scaffold19       AAATCAGATGTATGGGAAATAAGTCGATTTACATTAGATCGTACATTTGAACGAAACTAC
palI                  AAATCAGATGTATGGGAAATAAGTCGATTTACATTAGATCGTACATTTGAACGAAACTAC
                      ************************************************************


g423_contig1          GATAATGGTGTAAGTAATATTACAAAAGAGATGTTTTCAGCTCTATACGAGTTTGCTATT
g412_scaffold19       GATAATGGTGTAAGTAATATTACAAAAGAGATGTTTTCAGCTCTATACGAGTTTGCTATT
palI                  GATAATGGTGTAAGTAATATTACAAAAGAGATGTTTTCAGCTCTATACGAGTTTGCTATT
                      ************************************************************


g423_contig1          GCAAATTCAATCAATAATTTTGTCCTTGTAACTACAGTATCCTGTGAACGAATATTACGA
g412_scaffold19       GCAAATTCAATCAATAATTTTGTCCTTGTAACTACAGTATCCTGTGAACGAATATTACGA
palI                  GCAAATTCAATCAATAATTTTGTCCTTGTAACTACAGTATCCTGTGAACGAATATTACGA
                      ************************************************************


g423_contig1          TTATTAGGTATTCCAACTCGACGCTTTGGGGATGGAATGTCAAGTAAACTAGGAAAAGTA
g412_scaffold19       TTATTAGGTATTCCAACTCGACGCTTTGGGGATGGAATGTCAAGTAAACTAGGAAAAGTA
palI                  TTATTAGGTATTCCAACTCGACGCTTTGGGGATGGAATGTCAAGTAAACTAGGAAAAGTA
                      ************************************************************
```

```
g423_contig1          GATTCTGTAGCATTATGGGTTGATATAAATGACTCTTATAAACAAGCCGTAAATTGCAAA
g412_scaffold19       GATTCTGTAGCATTATGGGTTGATATAAATGACTCTTATAAACAAGCCGTAAATTGCAAA
palI                  GATTCTGTAGCATTATGGGTTGATATAAATGACTCTTATAAACAAGCCGTAAATTGCAAA
                      ************************************************************

g423_contig1          AAAATCAGTAATAATCAATATCTTATTGGTTGTTTATAAAAATTTGGACTGAGCCCAAGA
g412_scaffold19       AAAATCAGTAATAATCAATATCTTATTGGTTGTTTATAAAAATTTGGACTGAGCCCAAGA
palI                  AAAATCAGTAATAATCAATATCTTATTGGTTGTTTATAA--------------------
                      ************************************

g423_contig1          TAATTTCACAAAAAATATATTAAGCATTCTTTATCGCTAATTTTATATATTTAGAATCGG
g412_scaffold19       TAATTTCACAAAAAATATATTAAGCATTCTTTATCGCTAATTTTATATATTTAGAATCGG
palI                  ------------------------------------------------------------
```

*** *palR1* gene and its promoter**

```
g423_contig1        GTCGTAATGAATTGTAATGAAGATGGTGGTCAAACGGTATATCACATCCACCTTCATATG
g412_scaffold1      GTCGTAATGAATTGTAATGAAGATGGTGGTCAAACGGTATATCACATCCACCTTCATATG
palR1               ------------------------------------------------------------

g423_contig1        CTTGCAGGTAAAGAAATGGGCTGGCCGCCATATACAAATAATAAAAAAGTATTAATTTAA
g412_scaffold1      CTTGCAGGTAAAGAAATGGGCTGGCCGCCATATACAAATAATAAAAAAGTATTAATTTAA
palR1               ------------------------------------------------------------

g423_contig1        AAATAAATTAAAATAAAAGCAGCAAAAGAAACCCTAAAGTATTAGTTGAAACAAAAAACT
g412_scaffold1      AAATAAATTAAAATAAAAGCAGCAAAAGAAACCCTAAAGTATTAGTTGAAACAAAAAACT
palR1               ------------------------------------------------------------

g423_contig1        AATACTAAGGGTGTTATTTCTATGTTAAAGTTAGCCCTAATTTGAAATTTAGATACCTAT
g412_scaffold1      AATACTAAGGGTGTTATTTCTATGTTAAAGTTAGCCCTAATTTGAAATTTAGATACCTAT
palR1               ------------------------------------------------------------

g423_contig1        TTGAAGACAACGAACAAATAGGTAAGGGTAATTAACATGAAAAGTACTGCTTTTATTCTA
g412_scaffold1      TTGAAGACAACGAACAAATAGGTAAGGGTAATTAACATGAAAAGTACTGCTTTTATTCTA
palR1               ----------------------------------ATGAAAAGTACTGCTTTTATTCTA
                                                       ************************

g423_contig1        CAAGAACCTAATCCAATTAATGATATAAGCTTAGATGTACTAGCGCCTTTATTAAAAACC
g412_scaffold1      CAAGAACCTAATCCAATTAATGATATAAGCTTAGATGTACTAGCGCCTTTATTAAAAACC
palR1               CAAGAACCTAATCCAATTAATGATATAAGCTTAGATGTACTAGCGCCTTTATTAAAAACC
                    ************************************************************

g423_contig1        CAAGGATTGGAAGTTAAAAGTAGTACAGATACATCTGATATCCCTACTAATACTCGACTG
g412_scaffold1      CAAGGATTGGAAGTTAAAAGTAGTACAGATACATCTGATATCCCTACTAATACTCGACTG
palR1               CAAGGATTGGAAGTTAAAAGTAGTACAGATACATCTGATATCCCTACTAATACTCGACTG
                    ************************************************************

g423_contig1        TTGTTTATTGAATCAGGTGAAGAAAATGCATGGGAGAAATTGCAAGCGAAAGTAGCTGCA
g412_scaffold1      TTGTTTATTGAATCAGGTGAAGAAAATGCATGGGAGAAATTGCAAGCGAAAGTAGCTGCA
palR1               TTGTTTATTGAATCAGGTGAAGAAAATGCATGGGAGAAATTGCAAGCGAAAGTAGCTGCA
                    ************************************************************

g423_contig1        CTTACTTTTGAATGTGATATCGTATTATTTAATATTAGCCAAAATACTGAGCTTGCAAAT
g412_scaffold1      CTTACTTTTGAATGTGATATCGTATTATTTAATATTAGCCAAAATACTGAGCTTGCAAAT
palR1               CTTACTTTTGAATGTGATATCGTATTATTTAATATTAGCCAAAATACTGAGCTTGCAAAT
                    ************************************************************

g423_contig1        CGTGCACTATTAAGTGGTATCCGAGGTGTATTTTATGCCAATGATAATGCTGATGTATTA
g412_scaffold1      CGTGCACTATTAAGTGGTATCCGAGGTGTATTTTATGCCAATGATAATGCTGATGTATTA
palR1               CGTGCACTATTAAGTGGTATCCGAGGTGTATTTTATGCCAATGATAATGCTGATGTATTA
                    ************************************************************

g423_contig1        ATGAAAGGTATCCGATTATTACTTGAAAACCAACTTTGGTACCGAAGAGATATTATGTGC
g412_scaffold1      ATGAAAGGTATCCGATTATTACTTGAAAACCAACTTTGGTACCGAAGAGATATTATGTGC
palR1               ATGAAAGGTATCCGATTATTACTTGAAAACCAACTTTGGTACCGAAGAGATATTATGTGC
                    ************************************************************

g423_contig1        AATGCCTTAACTCGGTTATTACACTTTAATAAGGAAACAATCGCTAAGTTTACAGATGCG
g412_scaffold1      AATGCCTTAACTCGGTTATTACACTTTAATAAGGAAACAATCGCTAAGTTTACAGATGCG
palR1               AATGCCTTAACTCGGTTATTACACTTTAATAAGGAAACAATCGCTAAGTTTACAGATGCG
                    ************************************************************
```

```
g423_contig1        CCTGTAGAACCAGTGAACTTAACAAAGCGTGAACGTGCAATCATAACCTTGATGAGTACA
g412_scaffold1      CCTGTAGAACCAGTGAACTTAACAAAGCGTGAACGTGCAATCATAACCTTGATGAGTACA
palR1               CCTGTAGAACCAGTGAACTTAACAAAGCGTGAACGTGCAATCATAACCTTGATGAGTACA
                    ************************************************************

g423_contig1        GGCAGTAAAAATAAAGAGATCGCCGACAAGTTAGATATCAGCCCTCATACAGTAAAAACA
g412_scaffold1      GGCAGTAAAAATAAAGAGATCGCCGACAAGTTAGATATCAGCCCTCATACAGTAAAAACA
palR1               GGCAGTAAAAATAAAGAGATCGCCGACAAGTTAGATATCAGCCCTCATACAGTAAAAACA
                    ************************************************************

g423_contig1        CACTTATATAGTGCTTTTAGAAAAACAAAATGCCGCAATAGAATCGAGCTATTATCATGG
g412_scaffold1      CACTTATATAGTGCTTTTAGAAAAACAAAATGCCGCAATAGAATCGAGCTATTATCATGG
palR1               CACTTATATAGTGCTTTTAGAAAAACAAAATGCCGCAATAGAATCGAGCTATTATCATGG
                    ************************************************************

g423_contig1        GCACAACATAACATTCCTAATGAATTAAGATAGTAATAAAATAAAAACCAGCTTTAGCTG
g412_scaffold1      GCACAACATAACATTCCTAATGAATTAAGATAGTAATAAAATAAAAACCAGCTTTAGCTG
palR1               GCACAACATAACATTCCTAATGAATTAAGATAG--------------------------
                    *********************************

g423_contig1        GTTTTTTATATCTACAACTTTATATAAAACCTTTGCCATACCAATTCTATTAAGCTAATT
g412_scaffold1      GTTTTTTATATCTACAACTTTATATAAAACCTTTGCCATACCAATTCTATTAAGCTAATT
palR1               ------------------------------------------------------------
```

*** *palR2* gene and its promoter**

```
g423_contig1        TTTAATGAAAGCTGAATATATTTTAAGTTTATTTAATTAAAAATAATTAGATGAATTTAT
g412_scaffold15     TTTAATGAAAGCTGAATATATTTTAAGTTTATTTAATTAAAAATAATTAGATGAATTTAT
palR2               ------------------------------------------------------------

g423_contig1        TAAATGAGAGCTTTTGAATGAATGTAATTTTTGTTTTTTTATTAAATGCTGAATTGATTAT
g412_scaffold15     TAAATGAGAGCTTTTGAATGAATGTAATTTTTGTTTTTTTATTAAATGCTGAATTGATTAT
palR2               ------------------------------------------------------------

g423_contig1        CTATTTTATGGGAGTGCGCTGATTTATACGAAAAATTGATTTAATTAAAGTATAGACACT
g412_scaffold15     CTATTTTATGGGAGTGCGCTGATTTATACGAAAAATTGATTTAATTAAAGTATAGACACT
palR2               ------------------------------------------------------------

g423_contig1        AAAACCCATAAGGGGTACATTTGTAATCACTGGATTACAATTAGGTTAGTGACAAATTAA
g412_scaffold15     AAAACCCATAAGGGGTACATTTGTAATCACTGGATTACAATTAGGTTAGTGACAAATTAA
palR2               ------------------------------------------------------------

g423_contig1        CAAAGGAAGCTTATGATCTCTACAGCAACAATAAATACAGCTTATATTGTTACATCAGAT
g412_scaffold15     CAAAGGAAGCTTATGATCTCTACAGCAACAATAAATACAGCTTATATTGTTACATCAGAT
palR2               ------------ATGATCTCTACAGCAACAATAAATACAGCTTATATTGTTACATCAGAT
                                 *********************************************

g423_contig1        GAGACTCTGCATAGCAGCTCTCATAAGGTAGCTTTAACTAGCTTTATCGGTTTAATAGCA
g412_scaffold15     GAGACTCTGCATAGCAGCTCTCATAAGGTAGCTTTAACTAGCTTTATCGGTTTAATAGCA
palR2               GAGACTCTGCATAGCAGCTCTCATAAGGTAGCTTTAACTAGCTTTATCGGTTTAATAGCA
                    ************************************************************

g423_contig1        ACCTGTTCAAATAAGGTCCTCAAACAGAATAAGAGTGTTGATAATCAGTTTAATAAAATT
g412_scaffold15     ACCTGTTCAAATAAGGTCCTCAAACAGAATAAGAGTGTTGATAATCAGTTTAATAAAATT
palR2               ACCTGTTCAAATAAGGTCCTCAAACAGAATAAGAGTGTTGATAATCAGTTTAATAAAATT
                    ************************************************************

g423_contig1        GATGGAATATATTTTATTGATTTATTTTATTGTAATTGGAACAACAAAATACCTGATGAT
g412_scaffold15     GATGGAATATATTTTATTGATTTATTTTATTGTAATTGGAACAACAAAATACCTGATGAT
palR2               GATGGAATATATTTTATTGATTTATTTTATTGTAATTGGAACAACAAAATACCTGATGAT
                    ************************************************************

g423_contig1        ATTTTACAGTTAGCACAAAGGTCAAAAATTGTTTTATTTAATGTCCAAAATGACCAGTTA
g412_scaffold15     ATTTTACAGTTAGCACAAAGGTCAAAAATTGTTTTATTTAATGTCCAAAATGACCAGTTA
palR2               ATTTTACAGTTAGCACAAAGGTCAAAAATTGTTTTATTTAATGTCCAAAATGACCAGTTA
                    ************************************************************

g423_contig1        TGCGAAAAAAATTTATTGTTAGCGGGTTTTGAAGGTATTTTTTACCTTTCAGATAGACCC
g412_scaffold15     TGCGAAAAAAATTTATTGTTAGCGGGTTTTGAAGGTATTTTTTACCTTTCAGATAGACCC
palR2               TGCGAAAAAAATTTATTGTTAGCGGGTTTTGAAGGTATTTTTTACCTTTCAGATAGACCC
                    ************************************************************

g423_contig1        GATCTTATATTAAGAGGGTTAAACCAAATTAAAAATAATGAACGTTGGTTTAAACGCTCA
g412_scaffold15     GATCTTATATTAAGAGGGTTAAACCAAATTAAAAATAATGAACGTTGGTTTAAACGCTCA
palR2               GATCTTATATTAAGAGGGTTAAACCAAATTAAAAATAATGAACGTTGGTTTAAACGCTCA
                    ************************************************************

g423_contig1        TCAATGAATAATGCTTTTTCATATTTACTAAAGTCAAATAAAGCTTCAATATCACCTTCC
g412_scaffold15     TCAATGAATAATGCTTTTTCATATTTACTAAAGTCAAATAAAGCTTCAATATCACCTTCC
palR2               TCAATGAATAATGCTTTTTCATATTTACTAAAGTCAAATAAAGCTTCAATATCACCTTCC
                    ************************************************************
```

```
g423_contig1        AATAGCATCACAGGAAAATCTGTATTTCCAACGCTTACTAAACGTGAAAATACAATTATT
g412_scaffold15     AATAGCATCACAGGAAAATCTGTATTTCCAACGCTTACTAAACGTGAAAATACAATTATT
palR2               AATAGCATCACAGGAAAATCTGTATTTCCAACGCTTACTAAACGTGAAAATACAATTATT
                    ************************************************************

g423_contig1        AAATTAGTGACAAAAGGGTCTCAAAATCAAGAAATAGCTGATCAACTAAATATAAGTACA
g412_scaffold15     AAATTAGTGACAAAAGGGTCTCAAAATCAAGAAATAGCTGATCAACTAAATATAAGTACA
palR2               AAATTAGTGACAAAAGGGTCTCAAAATCAAGAAATAGCTGATCAACTAAATATAAGTACA
                    ************************************************************

g423_contig1        AATACGGTTAAAACCCATATTTATAGTATTTTTAGAAAGACTAAATCACGAAATCGTATT
g412_scaffold15     AATACGGTTAAAACCCATATTTATAGTATTTTTAGAAAGACTAAATCACGAAATCGTATT
palR2               AATACGGTTAAAACCCATATTTATAGTATTTTTAGAAAGACTAAATCACGAAATCGTATT
                    ************************************************************

g423_contig1        GAATTAATTACTTGGTCTTTACAATCATCAGGCCATTTAGATGCAGCCATTAATTAAATT
g412_scaffold15     GAATTAATTACTTGGTCTTTACAATCATCAGGCCATTTAGATGCAGCCATTAATTAAATT
palR2               GAATTAATTACTTGGTCTTTACAATCATCAGGCCATTTAGATGCAGCCATTAATTAA---
                    ************************************************************

g423_contig1        TTAAAGTAGCAACTAAATGCTACTTTTAATCTCGCTTATTTATATTGGTTATGATGTAAA
g412_scaffold15     TTAAAGTAGCAACTAAATGCTACTTTTAATCTCGCTTATTTATATTGGTTATGATGTAAA
palR2               ------------------------------------------------------------
```

*** *palR3* gene and its promoter**

```
g423_contig1        TCTAAAGCGCTTTTAACTCCCGTTGAATGGGAAGTTAATTAGTACAATTGGCATAACTTA
g412_scaffold19     TCTAAAGCGCTTTTAACTCCCGTTGAATGGGAAGTTAATTAGTACAATTGGCATAACTTA
palR3               ------------------------------------------------------------

g423_contig1        AATTAATAAATCACATCTTGCTTATCTAAAACTTCTATAAAATTAGAATACCTTTTTTTT
g412_scaffold19     AATTAATAAATCACATCTTGCTTATCTAAAACTTCTATAAAATTAGAATACCTTTTTTTT
palR3               ------------------------------------------------------------

g423_contig1        GAATTAATATTTTATATAGAATCAATAAATTATAAATATTGTATTAGTAATTAAAATGTT
g412_scaffold19     GAATTAATATTTTATATAGAATCAATAAATTATAAATATTGTATTAGTAATTAAAATGTT
palR3               ------------------------------------------------------------

g423_contig1        ATTTAAACGTAAATAACACTGTTAAGAAAATCGTTTTTAGGTACAATAATGATAGCGCTG
g412_scaffold19     ATTTAAACGTAAATAACACTGTTAAGAAAATCGTTTTTAGGTACAATAATGATAGCGCTG
palR3               ------------------------------------------------------------

g423_contig1        TCATTTCTGTGGCGCTAGTCTTATATTGAAGATAATAAAAGGGTTATATAAAATAATGCA
g412_scaffold19     TCATTTCTGTGGCGCTAGTCTTATATTGAAGATAATAAAAGGGTTATATAAAATAATGCA
palR3               ------------------------------------------------------ATGCA
                                                                           *****

g423_contig1        GTACAGTTGGTTAAATGGAATTATTGAAGGAATTAAACAAAGTAAAAATATAGACGATAT
g412_scaffold19     GTACAGTTGGTTAAATGGAATTATTGAAGGAATTAAACAAAGTAAAAATATAGACGATAT
palR3               GTACAGTTGGTTAAATGGAATTATTGAAGGAATTAAACAAAGTAAAAATATAGACGATAT
                    ************************************************************

g423_contig1        AAAATCTCAATGTGAAGCAATATGTCGTGCATTAGAAATTGATTTTTATTCGTTTGTAAT
g412_scaffold19     AAAATCTCAATGTGAAGCAATATGTCGTGCATTAGAAATTGATTTTTATTCGTTTGTAAT
palR3               AAAATCTCAATGTGAAGCAATATGTCGTGCATTAGAAATTGATTTTTATTCGTTTGTAAT
                    ************************************************************

g423_contig1        TAGGATACCAAGTTCACTTTTTAGCCCCGAAATTATTACATTATCTAATTACCCTCAATT
g412_scaffold19     TAGGATACCAAGTTCACTTTTTAGCCCCGAAATTATTACATTATCTAATTACCCTCAATT
palR3               TAGGATACCAAGTTCACTTTTTAGCCCCGAAATTATTACATTATCTAATTACCCTCAATT
                    ************************************************************

g423_contig1        ATGGCAAGAACATTATTTTTCTCAAGAGTTTATGAGTTTAGACCCTGTTATAGCGTATAG
g412_scaffold19     ATGGCAAGAACATTATTTTTCTCAAGAGTTTATGAGTTTAGACCCTGTTATAGCGTATAG
palR3               ATGGCAAGAACATTATTTTTCTCAAGAGTTTATGAGTTTAGACCCTGTTATAGCGTATAG
                    ************************************************************

g423_contig1        TCAACAACATATAACTCCAGTAAATTGGAGTGACTTAACTAACTATACTGACTTCAATTC
g412_scaffold19     TCAACAACATATAACTCCAGTAAATTGGAGTGACTTAACTAACTATACTGACTTCAATTC
palR3               TCAACAACATATAACTCCAGTAAATTGGAGTGACTTAACTAACTATACTGACTTCAATTC
                    ************************************************************

g423_contig1        CCCAAAAAAAATTAACGGTGTTAAAAGAGGCAAAAACTTATGGGTTATGCACTGGGGTCAG
g412_scaffold19     CCCAAAAAAAATTAACGGTGTTAAAAGAGGCAAAAACTTATGGGTTATGCACTGGGGTCAG
palR3               CCCAAAAAAAATTAACGGTGTTAAAAGAGGCAAAAACTTATGGGTTATGCACTGGGGTCAG
                    ************************************************************

g423_contig1        TATTCCACTTAAAGCTCCAACAGGTGAAATGAGTGTTTTTAGTTTGTCTTCGACGTCTCA
g412_scaffold19     TATTCCACTTAAAGCTCCAACAGGTGAAATGAGTGTTTTTAGTTTGTCTTCGACGTCTCA
palR3               TATTCCACTTAAAGCTCCAACAGGTGAAATGAGTGTTTTTAGTTTGTCTTCGACGTCTCA
                    ************************************************************
```

```
g423_contig1        AAGGTGTGTGATAACTGAGTGCCAGATAAATATACAATTACAAGCGCAAGTGATTGCTCC
g412_scaffold19     AAGGTGTGTGATAACTGAGTGCCAGATAAATATACAATTACAAGCGCAAGTGATTGCTCC
palR3               AAGGTGTGTGATAACTGAGTGCCAGATAAATATACAATTACAAGCGCAAGTGATTGCTCC
                    ************************************************************

g423_contig1        ATATCTTCATGAAGCCATTAAAATAATTAATTATAAGGCAAAAGAGATACTGAGCCATGA
g412_scaffold19     ATATCTTCATGAAGCCATTAAAATAATTAATTATAAGGCAAAAGAGATACTGAGCCATGA
palR3               ATATCTTCATGAAGCCATTAAAATAATTAATTATAAGGCAAAAGAGATACTGAGCCATGA
                    ************************************************************

g423_contig1        TGAGGTTAAAATAACTAATAGGGAAGAGGAGTGCTTATTGTGGGCCTGTGAAGGGAAAAC
g412_scaffold19     TGAGGTTAAAATAACTAATAGGGAAGAGGAGTGCTTATTGTGGGCCTGTGAAGGGAAAAC
palR3               TGAGGTTAAAATAACTAATAGGGAAGAGGAGTGCTTATTGTGGGCCTGTGAAGGGAAAAC
                    ************************************************************

g423_contig1        AAGCTGGGAGATTTCTAAAATAATAGGTATATCTGAAAGAACTGTGCTTTTTCATTTAAA
g412_scaffold19     AAGCTGGGAGATTTCTAAAATAATAGGTATATCTGAAAGAACTGTGCTTTTTCATTTAAA
palR3               AAGCTGGGAGATTTCTAAAATAATAGGTATATCTGAAAGAACTGTGCTTTTTCATTTAAA
                    ************************************************************

g423_contig1        TAATGTAAGCCAAAAAGTAGGGGGTGTAAATAGGCAACACAGTGTAGCTAAAGCCCTTTT
g412_scaffold19     TAATGTAAGCCAAAAAGTAGGGGGTGTAAATAGGCAACACAGTGTAGCTAAAGCCCTTTT
palR3               TAATGTAAGCCAAAAAGTAGGGGGTGTAAATAGGCAACACAGTGTAGCTAAAGCCCTTTT
                    ************************************************************

g423_contig1        AAATGGACTTATTCAGCCTAAGTTTTAATGAATTGTTATCGTTTTATTTAAAATACCTTT
g412_scaffold19     AAATGGACTTATTCAGCCTAAGTTTTAATGAATTGTTATCGTTTTATTTAAAATACCTTT
palR3               AAATGGACTTATTCAGCCTAAGTTTTAA------------------------------
                    ****************************
```

*** *palR4* gene and its promoter**

```
g423_contig1          CAGATAATGTGACTTTTATGTGATTTTCATAAAAGTCATATCGTTATTATTGTTTTATGC
g412_scaffold51_Rv    CAGATAATGTGACTTTTATGTGATTTTCATAAAAGTCATATCGTTATTATTGTTTTATGC
palR4                 ------------------------------------------------------------

g423_contig1          TTTTATTTATTTAAAGTTATTCCTGTTTTTTATTAAATATTTATAGGGTGCAACAAATAAT
g412_scaffold51_Rv    TTTTATTTATTTAAAGTTATTCCTGTTTTTTATTAAATATTTATAGGGTGCAACAAATAAT
palR4                 ------------------------------------------------------------

g423_contig1          GAAATCATTAAATTTATTAATAATTTTAACTGCAGTTTTTACAATGAACTTAGTTTTACT
g412_scaffold51_Rv    GAAATCATTAAATTTATTAATAATTTTAACTGCAGTTTTTACAATGAACTTAGTTTTACT
palR4                 ------------------------------------------------------------

g423_contig1          AGATTCTATTTTAATCTCTGAAACAGATTCTATTTCTCCTTTTAGTGTTTGTAAGATATT
g412_scaffold51_Rv    AGATTCTATTTTAATCTCTGAAACAGATTCTATTTCTCCTTTTAGTGTTTGTAAGATATT
palR4                 ------------------------------------------------------------

g423_contig1          AGATACATGCAAAAAGGAAGGGTAATTCTTAAAATTTGTGAGTTGGGGTTAACATGCAGT
g412_scaffold51_Rv    AGATACATGCAAAAAGGAAGGGTAATTCTTAAAATTTGTGAGTTGGGGTTAACATGCAGT
palR4                 -----------------------------------------------------ATGCAGT
                                                                            *******

g423_contig1          ACATCGAAAGCTTAATTAAAGAAATTAACTGTATTGATTCTAAAGAGCAGTTAGTTTTAT
g412_scaffold51_Rv    ACATCGAAAGCTTAATTAAAGAAATTAACTGTATTGATTCTAAAGAGCAGTTAGTTTTAT
palR4                 ACATCGAAAGCTTAATTAAAGAAATTAACTGTATTGATTCTAAAGAGCAGTTAGTTTTAT
                      ************************************************************

g423_contig1          TCTTTGAAAAAGTAACTGATTTATTTAATTATAAATGGTCTGGTTTAGTCATTTTCAAAC
g412_scaffold51_Rv    TCTTTGAAAAAGTAACTGATTTATTTAATTATAAATGGTCTGGTTTAGTCATTTTCAAAC
palR4                 TCTTTGAAAAAGTAACTGATTTATTTAATTATAAATGGTCTGGTTTAGTCATTTTCAAAC
                      ************************************************************

g423_contig1          CAAAATTGAATAATAAATATGAAGTGAGTGTTCTTGGTAATATACCTATAAACATGCAAG
g412_scaffold51_Rv    CAAAATTGAATAATAAATATGAAGTGAGTGTTCTTGGTAATATACCTATAAACATGCAAG
palR4                 CAAAATTGAATAATAAATATGAAGTGAGTGTTCTTGGTAATATACCTATAAACATGCAAG
                      ************************************************************

g423_contig1          AAAGAATAAAAAAAAAGCATAGATATAAGTGATTATTGTTTAAATGAGATAGAACCAAAAT
g412_scaffold51_Rv    AAAGAATAAAAAAAAAGCATAGATATAAGTGATTATTGTTTAAATGAGATAGAACCAAAAT
palR4                 AAAGAATAAAAAAAAAGCATAGATATAAGTGATTATTGTTTAAATGAGATAGAACCAAAAT
                      ************************************************************

g423_contig1          CTTTAGTTATTAAAGGTAAAGCATCAATTGAAAGTAATAATCTTGAAATTTTGGTTATAC
g412_scaffold51_Rv    CTTTAGTTATTAAAGGTAAAGCATCAATTGAAAGTAATAATCTTGAAATTTTGGTTATAC
palR4                 CTTTAGTTATTAAAGGTAAAGCATCAATTGAAAGTAATAATCTTGAAATTTTGGTTATAC
                      ************************************************************

g423_contig1          CTATTAACGGTGTGGGATCTGAATTTGCTTGTTTAACTTTTTTTATTAAACTCAGAAAAAC
g412_scaffold51_Rv    CTATTAACGGTGTGGGATCTGAATTTGCTTGTTTAACTTTTTTTATTAAACTCAGAAAAAC
palR4                 CTATTAACGGTGTGGGATCTGAATTTGCTTGTTTAACTTTTTTTATTAAACTCAGAAAAAC
                      ************************************************************

g423_contig1          AAAGAGGGTTGGTAGTTGAAAAAATAGGCTGGTTTTGGCTAATGTTATCTTCATTTATCT
g412_scaffold51_Rv    AAAGAGGGTTGGTAGTTGAAAAAATAGGCTGGTTTTGGCTAATGTTATCTTCATTTATCT
palR4                 AAAGAGGGTTGGTAGTTGAAAAAATAGGCTGGTTTTGGCTAATGTTATCTTCATTTATCT
                      ************************************************************
```

```
g423_contig1            ATAATAAATACAAAAAAGAAATAGCAGATGATAGTCATAAAATGACTAAAAGAGAATTGG
g412_scaffold51_Rv      ATAATAAATACAAAAAAGAAATAGCAGATGATAGTCATAAAATGACTAAAAGAGAATTGG
palR4                   ATAATAAATACAAAAAAGAAATAGCAGATGATAGTCATAAAATGACTAAAAGAGAATTGG
                        ************************************************************

g423_contig1            AATGTATTAAATGGGCCTCAGATGGCAAAACTTCATGGGAAATTAGTCAGTTATTGTCTA
g412_scaffold51_Rv      AATGTATTAAATGGGCCTCAGATGGCAAAACTTCATGGGAAATTAGTCAGTTATTGTCTA
palR4                   AATGTATTAAATGGGCCTCAGATGGCAAAACTTCATGGGAAATTAGTCAGTTATTGTCTA
                        ************************************************************

g423_contig1            TCAGTCAAAGAACTGTTGATTTTCATTTAGCTAACTGCATAGTAAAAACAGATAGTATTA
g412_scaffold51_Rv      TCAGTCAAAGAACTGTTGATTTTCATTTAGCTAACTGCATAGTAAAAACAGATAGTATTA
palR4                   TCAGTCAAAGAACTGTTGATTTTCATTTAGCTAACTGCATAGTAAAAACAGATAGTATTA
                        ************************************************************

g423_contig1            ACCGACAGCAAGCAATTGTTAAATGTGCTTTAAATGGTCATTTATTAGTATAGCTAAAGT
g412_scaffold51_Rv      ACCGACAGCAAGCAATTGTTAAATGTGCTTTAAATGGTCATTTATTAGTATAGCTAAAGT
palR4                   ACCGACAGCAAGCAATTGTTAAATGTGCTTTAAATGGTCATTTATTAGTATAG-------
                        ***************************************************

g423_contig1            GCATTAAGCTAAAGCAAGTAATTGTTTTAGCTTTATTATTTTATGATAACTATGCATTAG
g412_scaffold51_Rv      GCATTAAGCTAAAGCAAGTAATTGTTTTAGCTTTATTATTTTATGATAACTATGCATTAG
palR4                   ------------------------------------------------------------
```

*** *palR5* gene and its promoter**

```
g412_scaffold56             GCCCAATTGCATCGCCAGATATGATCCAATGGTCACCTGGTTTACCAAAAACCCGCTCAG
contig1_1264000_1267000     GCCCAATTGCATCGCCAGATATGATCCAATGGTCACCTGGTTTACCAAAAACCCGCTCAG
palR5                       ------------------------------------------------------------

g412_scaffold56             GTAAGATTATGCGTCGTATTTTACGTAAAATAGCTGCAAATGAACATCAACAATTAGGTG
contig1_1264000_1267000     GTAAGATTATGCGTCGTATTTTACGTAAAATAGCTGCAAATGAACATCAACAATTAGGTG
palR5                       ------------------------------------------------------------

g412_scaffold56             ATACTTCAACACTTGCAGATCCTACGGTAGTTGAAGAGTTAATTGAAAACCGTCTTAATC
contig1_1264000_1267000     ATACTTCAACACTTGCAGATCCTACGGTAGTTGAAGAGTTAATTGAAAACCGTCTTAATC
palR5                       ------------------------------------------------------------

g412_scaffold56             GTTAATAGCTTGCTTAAGTATTAATTAAATGAGTAGTATAATGGCTGTTGTCTAACTGAG
contig1_1264000_1267000     GTTAATAGCTTGCTTAAGTATTAATTAAATGAGTAGTATAATGGCTGTTGTCTAACTGAG
palR5                       ------------------------------------------------------------

g412_scaffold56             ATAACAGCCATTTTTTTAGGAGTAAACTATGAGCAAGTTTTTGATAGCGGATGATCACCC
contig1_1264000_1267000     ATAACAGCCATTTTTTTAGGAGTAAACTATGAGCAAGTTTTTGATAGCGGATGATCACCC
palR5                       ----------------------------ATGAGCAAGTTTTTGATAGCGGATGATCACCC
                                                        ********************************

g412_scaffold56             TTTATTTCGTGAAGCATTAAAAGGAGCGCTTCAGAACGCTTTTAGTGAGCTGGCTGTATT
contig1_1264000_1267000     TTTATTTCGTGAAGCATTAAAAGGAGCGCTTCAGAACGCTTTTAGTGAGCTGGCTGTATT
palR5                       TTTATTTCGTGAAGCATTAAAAGGAGCGCTTCAGAACGCTTTTAGTGAGCTGGCTGTATT
                            ************************************************************

g412_scaffold56             TGAGTCCGATAACTTCAAATCCACCCTTGAAATTTTAGCAAAAGAAGATGATCTAGATAT
contig1_1264000_1267000     TGAGTCCGATAACTTCAAATCCACCCTTGAAATTTTAGCAAAAGAAGATGATCTAGATAT
palR5                       TGAGTCCGATAACTTCAAATCCACCCTTGAAATTTTAGCAAAAGAAGATGATCTAGATAT
                            ************************************************************

g412_scaffold56             CTTATTGTTAGATCTACATATGCCGGGTAATGATGATTTATATGGCTTAATTAGGATCCG
contig1_1264000_1267000     CTTATTGTTAGATCTACATATGCCGGGTAATGATGATTTATATGGCTTAATTAGGATCCG
palR5                       CTTATTGTTAGATCTACATATGCCGGGTAATGATGATTTATATGGCTTAATTAGGATCCG
                            ************************************************************

g412_scaffold56             TGAAGATCACCCAGAACTGCCTATTGCAGTGGTATCAGGCAGTGAAGAAATCAGTGTTGT
contig1_1264000_1267000     TGAAGATCACCCAGAACTGCCTATTGCAGTGGTATCAGGCAGTGAAGAAATCAGTGTTGT
palR5                       TGAAGATCACCCAGAACTGCCTATTGCAGTGGTATCAGGCAGTGAAGAAATCAGTGTTGT
                            ************************************************************

g412_scaffold56             ATCTAAAGTGATGGCATATGGCGCATTGGGTTTTATCCCTAAATCCTTATCATCGGTAGA
contig1_1264000_1267000     ATCTAAAGTGATGGCATATGGCGCATTGGGTTTTATCCCTAAATCCTTATCATCGGTAGA
palR5                       ATCTAAAGTGATGGCATATGGCGCATTGGGTTTTATCCCTAAATCCTTATCATCGGTAGA
                            ************************************************************

g412_scaffold56             GATTGCTGTTGCAATTAATGAAATATTAGAAGGCGAAACTTGGTTACCTGAAACAATGAA
contig1_1264000_1267000     GATTGCTGTTGCAATTAATGAAATATTAGAAGGCGAAACTTGGTTACCTGAAACAATGAA
palR5                       GATTGCTGTTGCAATTAATGAAATATTAGAAGGCGAAACTTGGTTACCTGAAACAATGAA
                            ************************************************************

g412_scaffold56             AGATAAGGTCAATCAATTATCAGGAGATGAAGTTAAAGTGGCGACACAAGTTGCATCTTT
contig1_1264000_1267000     AGATAAGGTCAATCAATTATCAGGAGATGAAGTTAAAGTGGCGACACAAGTTGCATCTTT
palR5                       AGATAAGGTCAATCAATTATCAGGAGATGAAGTTAAAGTGGCGACACAAGTTGCATCTTT
                            ************************************************************
```

```
g412_scaffold56             AACACCACAACAATATAAAGTGTTGAGTTATTTGCATGAAGGTCTATTGAATAAACAGAT
contig1_1264000_1267000     AACACCACAACAATATAAAGTGTTGAGTTATTTGCATGAAGGTCTATTGAATAAACAGAT
palR5                       AACACCACAACAATATAAAGTGTTGAGTTATTTGCATGAAGGTCTATTGAATAAACAGAT
                            ************************************************************

g412_scaffold56             TGCATATGAATTAAATATCTCTGAAGCAACGGTTAAAGCGCATATTACTGCTATTTTCAG
contig1_1264000_1267000     TGCATATGAATTAAATATCTCTGAAGCAACGGTTAAAGCGCATATTACTGCTATTTTCAG
palR5                       TGCATATGAATTAAATATCTCTGAAGCAACGGTTAAAGCGCATATTACTGCTATTTTCAG
                            ************************************************************

g412_scaffold56             AAAGTTAGGTGTTTATAATCGCACTCAAGCGGTATTGATTGCCTCTAAGTTACAATTAGA
contig1_1264000_1267000     AAAGTTAGGTGTTTATAATCGCACTCAAGCGGTATTGATTGCCTCTAAGTTACAATTAGA
palR5                       AAAGTTAGGTGTTTATAATCGCACTCAAGCGGTATTGATTGCCTCTAAGTTACAATTAGA
                            ************************************************************

g412_scaffold56             ATCGCCAGTAGAGGCTTAAAAATGTGCTAAAGTTTAATTATCCCCGAACTGTAATCGGGG
contig1_1264000_1267000     ATCGCCAGTAGAGGCTTAAAAATGTGCTAAAGTTTAATTATCCCCGAACTGTAATCGGGG
palR5                       ATCGCCAGTAGAGGCTTAA----------------------------------------
                            ******************

g412_scaffold56             ATATGTTTATTACATTCTCTTACTTAAACACTACCGTTTTATTACCATTAACAAACAGTC
contig1_1264000_1267000     ATATGTTTATTACATTCTCTTACTTAAACACTACCGTTTTATTACCATTAACAAACAGTC
palR5                       ------------------------------------------------------------
```

REAGENTS PREPARATION

Name of medium, reagent		*Composition*	
LB medium (pH 7.0)		Tryptone	10.0 g/L
		Yeast extract	5.0 g/L
		NaCl	10.0 g/L
		Agar	15.0 g/L
PPES-II medium (pH 7.8)		Polypeptone	2.0 g/L
		Proteose Peptone No.3	1.0 g/L
		Soytone Peptone	1.0 g/L
		Yeast Extract	1.0 g/L
		Ferric Citrate	0.1 g/L
AB minimal medium * **S1:** 50 ml **S2:** 50 ml **S3:** 10 ml **H_2O:** 890 ml *: S1, S2, S3 are autoclaved separately	**Solution 1 (S1)**	K_2HPO_4	60.0 g/L
		NaH_2PO_4	20.0 g/L
	Solution 2 (S2)	NH_4Cl	20.0 g/L
		$MgSO_4 \cdot 7H_2O$	6.0 g/L
		KCl	3.0 g/L
		$CaCl_2 \cdot 2H_2O$	0.2 g/L
		$FeSO_4 \cdot 7H_2O$	0.05 g/L
	Glucose (S3)	Glucose	20%
Ampicillin stock solution 100 mg/ml (filter-sterilized, stored at -20 ^oC)		Ampicillin	5 g
		Distilled water	to 50 ml
Kanamycin stock solution 50 mg/ml (filter-sterilized, stored at -20 ^oC)		Kanamycin sulfate	2.5 g
		Distilled water	to 50 ml
Gentamicin stock solution 5 mg/ml (filter-sterilized, stored at -20 ^oC)		Gentamicin sulfate	5.0 mg
		Distilled water	to 1 ml