Norwegian University of Life Sciences
Faculty of Veterinary Medicine and Biosciences
Department of Chemistry, Biotechnology and
Food Science

Master Thesis 2014
60 credits

Identification of RIP metallopeptidase RseP as Target Receptor for the Leaderless Bacteriocin Ej97, and the Presumptive Involvement of the Ecs ABC transporter in medium Resistance

Marianne Slang Jensen

Acknowledgements

The work implemented in this thesis was performed throughout 2014 at the Laboratory of Microbial Gene Technology (LMG) at the Norwegian University of Life Sciences (NMBU), with Professor Dzung Bao Diep and Professor Helge Holo as supervisors.

I would like to begin by thanking Dzung for accepting me as his Masters student, and for being ever enthusiastic and inspiring in his guidance. You have always met me with genuine interest, some bad jokes, and invigorating thoughts and discussions in abundance.

Several lab members deserve special thanks for helping me complete my thesis. This project is a continuation of the work of Kirill Ovchinnikov, who throughout this year has indulged me with interesting perspectives, suggestions, and a bit of criticism when appropriate. I would also like to thank Ibrahim Mehmeti who helped me get my project off to the best possible start. Özgün Candan Onarman Umu, thank you for all your patience with me in the preparations of whole genome sequencing. You are a perfectionist far surpassing me and I could not have wished for a better teacher. Cyril Frantzen, I am grateful for all your help with the MiSeq, I could not have done it without you and would quite possibly still be stuck at this point. Thank you for devoting your time to my silly beginner's questions and guiding me through the analysis process.

Thanks to all the people at LMG for a great year which I have sincerely appreciated more than I could have imagined. Little could go wrong with Linda Godager ensuring the lab was always in order. May-Britt Selvåg Hovet, thank you for tirelessly wielding the autoclave, but also for making me feel right at home in the lab. Hai Chi, Eirik, Juan José, Andreza, Anne Kristin and Pawel, thank you all for good advice, for social events, and for a great working environment. I will be sad to leave you.

Finally I would like to thank my family and friends for all their support and encouragement.

Ås, December 2014

Marianne Slang Jensen

Abstract

In a recent study on the leaderless and *Lactococcus*-specific bacteriocin LsbB, the results implied that LsbB utilizes a zinc-dependent metallopeptidase YvjB from the M50 superfamily as receptor for target recognition and subsequent killing of other lactococcal bacteria.

Based on this evidence, another study was initiated which elucidated a conserved motif near the C-terminal of LsbB, proving to be crucial for its activity. The work also revealed that several other leaderless bacteriocins, including Ej97, K1 and EntQ, all share this conserved motif with LsbB, suggesting that they target the same receptor. Despite this similarity, these bacteriocins display a significantly broader inhibition spectrum than LsbB.

For this thesis, it was initially affirmed that Ej97 and K1 exhibit a wide inhibition spectrum relative to LsbB, followed by generation of approximately 80 spontaneous mutants resistant to Ej97, K1 or EntQ. These mutants belonged to pediococcal, enterococcal and lactococcal strains. The homologous genes of the LsbB receptor (*rseP*) were identified in order to be sequenced and analyzed in all the collected mutants. Severe mutations inflicting premature termination of RseP were in many cases revealed. This suggests that these RseP proteins which are closely related to the LsbB receptor also serve as docking molecules in other bacterial species for LsbB-like bacteriocins. They are however unable to function as receptor for LsbB.

Still, a large proportion of the mutants did not hold mutations within the putative RseP receptor. Phenotypic microtiter assays were performed to investigate Ej97 resistance levels. It became evident that the degree of resistance generally diverged into two levels – one very high indicating absolute resistance, and a lower level but still significantly higher than in the wild type. Interestingly, the mutants with the highest resistance level coincided with the presence of dysfunctional RseP.

Concluding the work for this thesis, whole genome sequencing of 40 mutants was carried out to identify possible reasons for the resistance observed in mutants with intact RseP and lower level of resistance. Variant detection provided clear indications of components from the Ecs ABC-transporter being involved. After this revelation it was noticed that previous studies on non-bacteriocin systems have indicated a connection between EcsAB and RseP in other species. Future work will examine how these genes are interrelated in conferring bacteriocin resistance.

Sammendrag

I et nylig studie på det lederløse og laktokokk-spesifikke bakteriosinet LsbB ble det konstantert at LsbB anvender en sink-avhengig metallopeptidase YvjB fra M50-superfamilien som reseptor for å gjenkjenne og angripe målceller.

Et påfølgende studie viste at en rekke andre lederløse bakteriosiner, inkludert Ej97, K1, og EntQ, deler et konservert motiv ved C-terminalen som er essensielt for aktiviteten hos LsbB. Dette tyder på at de benytter seg av den samme reseptoren, men til tross for denne likheten har disse slektningene betydelig bredere inhibisjonsspektra sammenlignet med LsbB.

I denne avhandlingen ble det i første omgang bekreftet at Ej97 og K1 har et bredt inhibisjonsspektrum i forhold til LsbB, og det ble deretter generert omkring 80 spontane mutanter resistente mot Ej97, K1 eller EntQ. Disse mutantene tilhørte pediokokke, enterokokke og laktokokke stammer. I hver av de aktuelle stammene ble det homologe genet til LsbB-reseptoren (*rseP*) identifisert. *rseP* kunne deretter bli amplifisert og sekvensert i samtlige mutanter, og det ble i mange tilfeller oppdaget alvorlige mutasjoner som medførte pre-terminering av RseP. Dette tydet på at disse RseP-membranproteinene som er nært beslektet med LsbB-reseptoren også kan fungere som reseptor i andre bakteriearter for LsbB-lignende bakteriosiner. Samtidig kan de derimot ikke fungere som reseptor for LsbB.

Imidlertid var det en stor andel av mutantene som ikke hadde mutasjoner innad det antatte reseptorgenet. Fenotypiske microtiter assays ble utført for å se på resistensgraden mot Ej97. Det viste seg at graden av resistens generelt kunne deles i to nivåer – ett svært høyt som indikerte full resistens, og et lavere men samtidig klart høyere resistensnivå enn hos villtypen. Mutantene med høyest resistens var de samme som hadde mutasjoner i *rseP*.

Avslutningsvis ble det foretatt helgenomsekvensering av 40 mutanter for å lokalisere mulige årsaker til resistensen man observerte i mutanter med inntakt RseP og lavere resistensnivå. Variantdeteksjon ga klare indikasjoner på at komponentene av Ecs ABC-transportøren er involvert. Etter denne oppdagelsen ble det bemerket at tidligere studier ikke relatert til bakteriosiner har avdekket sammenhenger mellom EcsAB og RseP i andre bakteriearter. Videre arbeid vil se nærmere på hvordan disse genene henger sammen i forhold til bakteriosinresistens.

Table of Contents

1. INTRODUCTION

1.1 Antimicrobial agents

Physical and chemical preservatives such as decreased pH, high salt concentration, low water content, and intense temperatures have been employed in foods for suppressing microbial growth for thousands of years. Stronger chemicals and treatments have additionally been exploited for decades in disinfection and sterilization to prevent crossover of contamination to foods and humans, but are generally not compatible with human consumption. These substances include components such as phenol, ethanol, chlorine, and peroxide, as well as newer compounds with prolonged activity. Among the utilized physical treatments are various heat- and radiation-based processes (APUA 2014; Todar 2008).

The first chemotherapeutic agents for fighting infectious diseases were certain alkaloids, i.e. plant extracts with strong physiological effects, and were put to use in the 1600s in South America primarily to treat malaria. Synthetic agents such as the arsenic-containing salvarsan was in the early 1900s developed to cure syphilis, but due to the toxicity it was not ideal for treatment. In the 1930s, the introduction of sulfonamides drastically improved treatment of infectious diseases, many of which are still being used today (Yazdankhah et al. 2013).

Bacterially produced substances with antagonistic activity comprise a large group, including traditional antibiotics, bacteriophages and antimicrobial peptides (AMPs), as well as various other inhibitory agents such as metabolic by-products (ammonia, hydrogen peroxide etc.) and bacteriolytic/autolytic enzymes (Jack et al. 1995; Tagg et al. 1976).

Traditional antibiotics have been utilized since the 1940s for targeting pathogens causing infectious diseases in humans and animals. Being discovered by Alexander Fleming in 1928 by accidental mold contamination on staphylococci, Howard Florey and Ernst Chain were responsible for putting the first antibiotic (penicillin) into large-scale production by 1942. The three scientists were later awarded the Nobel Prize in Physiology for their achievements. The new “miracle cure” was soon distributed to the public, proving especially valuable to soldiers during World War II in the following years. However, the limitations of penicillin were soon

uncovered, with problems concerning both the activity range as well as causing allergic reactions in some patients. Large sums of money were subsequently invested in discovering new types of antibiotics, initiating the antibiotic golden era. However, in the 1960s the development came to a sudden halt (Yazdankhah et al. 2013).

Since then, only one new antibiotic class (the oxazolidinones) have been discovered, and the lack of development within the field of traditional antibiotics for the past decades is not an ideal situation for maintaining effective treatment options (Hassan et al. 2012). An increasing problem concerning traditional antibiotics is the emergence of resistant and multi-drug-resistant (MDR) bacteria in response to prolonged and inadequate use. Poor practice is widespread in human healthcare due to overprescribing, but moreover in animal feeds for controlling disease and promote increased growth in large populations. This poses a threat of resistance crossing from animal products to humans. Supplementation of antibiotics in feed has thus been banned in several countries to restrict dispersion of resistance and MDR development, but the prevalence of resistant bacteria have already reached a level of great concern (Cheng et al. 2014).

Resistance develops and spreads by genetic mutations rendering a bacterium unsusceptible to the actions of the antibiotic, and the resistance-determinants subsequently being transferred between bacteria. This process is assisted by horizontal gene transfer through conjugation (transfer of genetic material between bacteria through cell-to-cell contact), transduction (facilitated by host DNA-uptake in virus) and transformation (incorporation of free DNA from the environment by competent cells). Together, these mechanisms contribute to restricting treatment options in patients and have severe consequences in regards to mortality rates and economic perspectives, e.g. the increasing occurrence of methicillin-resistant *Staphylococcus aureus* (MRSA) infections in hospitals and nursing homes (Huddleston 2014).

Many alternatives to traditional antibiotics have been proposed and are presently being investigated for various applications in animal feed, food preservation, and the pharmaceutical industry. They include antibacterial vaccines, immunomodulatory agents, bacteriophages and their lysins, probiotics, plant extracts, inhibitors of bacterial quorum sensing/biofilm/virulence, feed enzymes, and antimicrobial peptides (AMPs) (Cheng et al. 2014).

1.2 Bacteriocins

For millennia, human populations across the world have benefitted from the fermentative metabolism of lactic acid bacteria (LAB) and the resulting pH decrease, in order to prevent rapid microbial spoilage of foods. In addition to this general inhibition mechanism, scientists of the past century have revealed that most LAB also possess additional properties constituting mechanisms of a significantly more specific nature. These mechanisms involve the production and secretion of small antimicrobial peptides, termed bacteriocins, presumed to attach to unique receptor molecules on the target in order to exert potent activity. The bacteriocins are not to be confused with the colicins and microcins, which although also proteinaceous substances are produced and mainly active against Gram-negative species. These substances and their possible applications will not be discussed in detail here (Eijsink et al. 2002; Nes et al. 1996).

Production of AMPs is a widespread trait within both the prokaryotic and eukaryotic domain, but while contributing to the innate immunity system of eukaryotic life forms for protection against invading organisms, AMPs produced by bacteria are primarily considered as a means of strategic warfare between competing bacteria within the same niche (Nes et al. 2007; Nissen-Meyer & Nes 1997). The abundant use of AMPs in defense mechanisms has likely been evolutionally successful due to ribosomal synthesis of peptides being a rapid response process compared to e.g. production of antibiotics, antibodies, and immune cells (Nissen-Meyer & Nes 1997).

In the broadest sense of the word, bacteriocins are defined as ribosomally produced peptides with antimicrobial activity. The bacteriocins should be regarded as antimicrobial compounds separate from the traditional antibiotics due to being gene-encoded, while traditional antibiotics are secondary metabolites produced by the assistance of multi-enzyme complexes (Nes et al. 2007). Additionally, in contrast to traditional antibiotics which usually act as enzyme inhibitors (causing damage to cell-wall synthesis, DNA synthesis, protein synthesis, or other metabolic pathways), bacteriocins are generally known to exert their activity on target cells by permeabilization of the cell membrane causing leakage and loss of ion gradient (Diep et al. 2007).

By observing the inhibitory effects caused by bacteria isolated from urine (presumably *E. coli*) on the pathogen *Bacillus anthraces* (causing anthrax), Pasteur and Joubert were likely the first to record such antagonistic interactions between bacteria (Tagg et al. 1976). Recognition of these

antimicrobial agents as secreted peptides, later termed colicins, was first demonstrated by Gratia in 1925 when he observed that the *E. coli* strain V produced a heat-stable substance that effectively inhibited the growth of another *E. coli* strain ϕ (Reeves 1965; Tagg et al. 1976). Discovery of several other colicins and improved knowledge of their characteristics ensued in the following years, and the more general term "bacteriocin" was appointed by Jacob et al in 1953 after it had been established that these substances could be produced by other bacteria than coliform members of the Gram-negative Enterobacteriaceae family (Jack et al. 1995).

The characterization of bacteriocins from Gram-positive LAB species was heavily based on prior knowledge of the colicins. This included being proteinaceous substances with a relatively narrow inhibition spectrum and intraspecies killing activity at low concentrations, and which depended on the presence of a specific receptor on the target. Additionally, they were known to be resistant to heat inactivation and produced along with substances which rendered the producer itself immune. Their proteinaceous character could be affirmed by observing inactivation of the substances by certain proteases, as well as trypsin treatment being capable of rescuing susceptible target cells after having been exposed to a lethal dose of bacteriocin. The latter experiment also suggested a two-step mechanism, with the bacteriolytic activity being initiated from the receptor site after the bacteriocin had already docked. This conclusion was assumed after observing that trypsin rescue was possible only if initiated within the first couple of minutes after bacteriocin exposure (Reeves 1965; Tagg et al. 1976).

New knowledge has revealed that the bacteriocins of the Gram-positive bacteria comprise a highly heterogeneous group with a generally lower molecular weight than the colicins and displaying a somewhat less narrow activity spectrum, in some cases being active even towards Gram-negative species (Jack et al. 1995; Tagg et al. 1976).

1.2.1 Applications and future prospects in food safety and pharmaceutical industries

Nisin is presently the sole example of a commercially utilized bacteriocin widely implemented in food preservation. The peptide has been known since 1928 as a product of *Lactococcus lactis* strains, and was structurally characterized already in 1971 (Jack et al. 1995). By then, nisin-producing strains had already been implemented in dairy production for a long time without occurrence of any adverse effects. In 1988, nisin was approved for use as a food additive in the

US food industry after the Food and Drug Administration allowed nisin produced by certain strains to be employed to prevent outgrowth of *Clostridium botulinum* spores and toxin formation in pasteurized cheese spreads and canned foods (FDA, 1988).

The target cells of bacteriocins are generally confined to a relatively narrow selection of closely related LAB, but frequently also include prominent Gram-positive food spoilage bacteria and human pathogens. The peptides (although pH-tolerant and heat-stable) are easily degraded in the body and have no effect on human cells due to their distinct differences in composition relative to bacterial cells. Like their producers, they are thus generally recognized as safe substances (GRAS-status) for human consumption and handling. Additionally, many of the LAB known to produce bacteriocins which inhibit the growth of food spoilage bacteria are often found naturally in our intestine or already widely implemented in the food industry. These are all exceedingly attractive traits, establishing the LAB bacteriocins as promising tools for future industrial exploitations and safe applications in food preservatives and pharmaceutics (Eijsink et al. 2002).

Additionally, as the bacteriocins are ribosomally synthesized from dedicated genes and often contain minimal post-translational modifications, they are interesting subjects in peptide engineering for enhancing traits. Engineering inducing various modifications, gene-shuffling and re-design can possibly enhance the activity, potency and stability of the peptide. For this to be successfully executed with rational design, it is necessary to have profound knowledge of the structure-function relationship. However, modern techniques allow scientists to do high-throughput screening of random mutants and hybrids, facilitating directed evolution on a large scale (Jack et al. 1995; Lohans & Vederas 2012; Nes et al. 1996).

The administration of possible commercial bacteriocins could be executed in a variety of ways, including as purified additives or produced *in situ* (e.g. in bacterial starter cultures or as probiotics). In many cases it could prove favorable to administer the peptides in combination with other preservatives and antimicrobials, or in combination with other bacteriocins. This combination-strategy could potentially enable more natural preservation of foods with less need of chemicals and physical treatments (Galvez et al. 2007). LAB bacteriocins are commonly not active against Gram-negative species due to the outer membrane forming a protective barrier. However, by using a combination of bacteriocin and a chelating agent such as EDTA in simultaneous treatment for weakening the outer membrane, there is a possibility of expanding

the applications of bacteriocins and increase their value in the food industry. Trials with nisin have provided encouraging results in regards to such opportunities (Stevens et al. 1991).

Regardless, bacteriocins need to be vigorously tested in clinically relevant settings to evaluate their real potential in pharmaceutics, farming, and food industry. Additionally, the efficacy of bacteriocin synthesis is influenced by several environmental factors, and therefore the optimal conditions must also be determined for individual bacteriocins to make them economically viable in industries. For commercial use to be feasible, methods for large-scale production and purification with high yield must be further developed and improved (Galvez et al. 2007).

Another valuable trait of the bacteriocins are their activity spectra which range from very narrow to fairly wide, and it has long been known that strains frequently produce more than one bacteriocin. In treatment of infectious diseases, it is often preferable to employ narrow-spectrum agents in order to prevent disrupting the balance of the commensal gut-community, a problem often associated with traditional broad-spectrum antibiotics and which assist the emergence of opportunistic pathogens, capable of taking advantage of the sudden lack of competition. However, when the illness is of unknown origin, the better choice is rather broad-spectrum agents, and many known bacteriocins hold this trait as well. Due to bacteriocins targeting unique sites on susceptible cells and conducting their mode of action as permeabilization of the membrane, the probability of cross-resistance to traditional antibiotics is generally low. Several bacteriocins have already displayed promising activity against antibiotic resistant strains such as vancomycin-resistant enterococci (VRE) and opportunistic *Staphylococcus aureus* (MRSA) (Hassan et al. 2012; Lohans & Vederas 2012).

Like traditional antibiotics, it is likely that problems with resistance will occur with bacteriocins if they are put into commercial use. However, scientific developments since the introduction of traditional antibiotics can assist todays scientists in preforming the necessary studies for fully understanding the resistance mechanisms and thereby minimize this problem (Cotter et al. 2013).

1.2.2 Biosynthesis and mode of action

Bacteriocin production is an energy consuming and thus often strictly regulated process. The peptides are generally produced from dedicated loci with leader peptides attached up until processing and secretion from the cell. It is generally believed that bacteriocins kill other bacteria

by pore-formation in the membrane after receptor recognition, while the producers themselves remain immune to the bacteriocin activity by expressing specific immunity proteins.

Bacteriocin loci: In general, loci dedicated to bacteriocin production are dependent on several elements in addition to the structural bacteriocin genes; including components for processing and secretion, regulation, and self-immunity. The structural bacteriocin genes are usually located upstream the corresponding immunity genes and co-regulated with these, as illustrated in figure 1.1 (Eijsink et al. 2002). Commonly, a minimum of four genes are essential for bacteriocin production and are usually found in close proximity within the same loci. The structural genes encode pre-bacteriocins which have yet to be cleaved and folded, the immunity genes provide protection for the producer, dedicated membrane-bound ABC transporters facilitate exportation, and an accessory protein with elusive function is also needed for secretion. Exceptions to this organization are the class IIb bacteriocins whose activity require the action of two different peptides (two structural genes), and the class I bacteriocins (lantibiotics) which need additional genes for catalyzing dehydration of selected amino acids and lanthionine ring formation, and serine proteases to cleave off the leader peptide (Nissen-Meyer & Nes 1997).

Figure 1.1: Organization of the class IIa bacteriocin sakacin A *sap* locus. *sapA* encodes the bacteriocin precursor, and *saiA* the cognate immunity protein. *orf4* produces a Sap-Ph pheromone precursor. *sapK* codes for a sensory histidine protein kinase, *sapR* for the response regulator, *sapT* for the ABC transporter, and *sapE* for the cognate accessory protein. IS is an insertion sequence element, and the triangles marked "i" and "ii" indicate regulated promotors (Diep et al. 2000).

Production: Genetic determinants for bacteriocin production are occasionally chromosomally located, but more commonly plasmid-bound or even transposon-associated. As mobile genetic elements they are capable of promoting transfer to other bacteria by horizontal gene transfer. The status of bacteriocin synthesis rely on the physiological conditions surrounding the producer, including growth media composition, temperature, incubation time, pH, and aeration. These factors must be empirically investigated to establish optimal conditions for triggering bacteriocin synthesis. Production can in some cases also be induced by the assistance of certain substances or irradiation with ultraviolet light (Jack et al. 1995; Tagg et al. 1976).

Processing and secretion: Most bacteriocins are produced as linear pre-peptides with a leader sequence extension at the N-terminal, commonly in the form of a 14-30 aa double-glycine leader. This extension is responsible for maintaining inactivity of the peptide whilst inside the producer cell. The leader peptide is cleaved off upon bacteriocin secretion by dedicated ABC-transporter machinery, or in some cases by a sec-dependent pathway. The accessory protein is also essential for successful externalization of the peptide, although its specific role has yet to be established. After cleavage and modifications, the peptide reaches mature and active state after having been transported to the extracellular environment and attained its three dimensional folded structure (Eijsink et al. 2002; Nes et al. 1996).

Self-immunity: Genes conferring self-immunity are crucial to protect the producer from being harmed by the bacteriocin. They are commonly co-regulated with the bacteriocin and hold a conserved position downstream of the structural bacteriocin genes, making them relatively easy to identify. Still, most mechanisms of self-immunity remain elusive (Kjos et al, 2009). Studies indicate that immunity is conferred in various manners, with the presently best studied models being nisin and class IIa bacteriocins. Immunity in nisin-producing strains is conferred by two separate mechanisms, namely a specialized ABC transporter which pumps nisin out of the membrane, and a dedicated immunity protein (LanI) which interacts with nisin on the extracellular side and prevents it from binding to the receptor (Draper et al. 2008). In the case of class IIa producers, the results suggest that the immunity proteins bind tightly to the intercellular part of the receptor and thereby prevents pore formation in the presence of bacteriocin (Diep et al. 2007; Hassan et al. 2012).

Regulation: As processing and secretion of bacteriocins is dependent on ABC transporters and their ATP-binding component for coupling ATP hydrolysis with transportation, the production is energy-demanding and often strictly regulated to avoid wasting resources. Some bacteriocins, especially those of Gram-positive species, are known to be regulated according to cell-density in a phenomena known as quorum-sensing. By using secretion of pheromones to communicate with members of the same strain, the bacteria can co-ordinate production of bacteriocin (Nes et al. 1996). This mechanism enables the bacterium to sense its own growth compared to competing bacteria, ensuring that full bacteriocin production is only initiated when it is useful to the cell, i.e. when the environment contains a certain density (quorum) of neighboring cells. Class I

bacteriocins utilize the bacteriocin itself as inducing factor and is thus auto-regulated by the bacteriocin. Regulation of class II bacteriocins by quorum-sensing is dictated by pheromone induction instead, with the pheromone being structurally similar to the bacteriocin but without the antimicrobial activity (Snyder & Worobo 2014). This quorum-sensing mechanism is frequently termed three-component regulatory system. The system consists of three co-transcribed genes which encode a bacteriocin-like peptide (pheromone), a specific histidine kinase (sensor protein), and a DNA-binding effector protein (response regulator). Activation of bacteriocin production commences with secretion of the bacteriocin-like pheromone, which binds to the histidine kinase sensor protein in the membrane. Bacteria of the same strain will all secrete this pheromone at low levels, and when being in sufficient numbers reaching a certain threshold level, a chain reaction of auto-phosphorylation will be initiated. The phosphoryl group will during this process be transferred to the response regulator, which in turn becomes capable of binding DNA and activate the regulated promotor of the bacteriocin locus (figure 1.2). This induces massive pheromone and bacteriocin production (Eijsink et al. 2002).

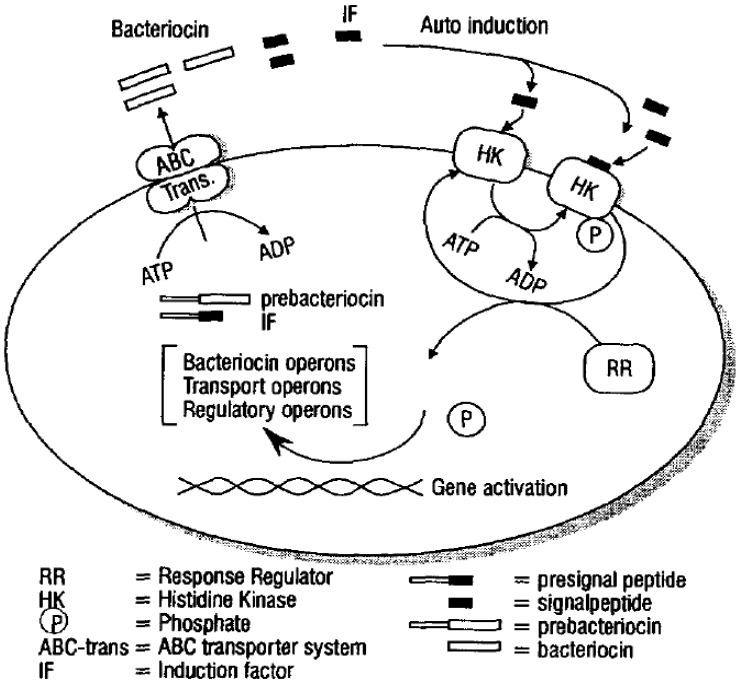


Figure 1.2: Schematic of how bacteriocin production can be regulated by quorum-sensing. The inducing factor IF (bacteriocin-like pheromone) binds to the N-terminal sensory domain of histidine kinase protein HK at the exterior of the membrane. This stimulates activation of the C-terminal autokinase activity of HK on the inside of the membrane. This activation initiates auto-phosphorylation, leading to the response regulator RR being phosphorylated and inducing transcription of bacteriocin-affiliated operons (Nes et al. 1996).

Initial interaction: Contact between extracellular bacteriocin and target cell membrane is facilitated by unspecific electrostatic interactions due to the positively charged cationic character of the bacteriocin and the negatively charged anionic lipids of the target membrane (Eijsink et al. 2002; Nissen-Meyer & Nes 1997). Evidence of the bacteriocins' usually amphiphilic structure suggested that the killing activity was enabled by the hydrophobic regions interacting with the target cell membrane and causing pore-formation (Nissen-Meyer & Nes 1997; Tagg et al. 1976).

Target recognition and mode of action: The specific interaction occurs when the bacteriocin recognizes and docks with a specific receptor on the surface of the target cell. The spectrum of activity is thus determined by which bacteria contain the receptor molecule compatible with the bacteriocin. The common mode of action is so far known to be attachment to the receptor, and subsequently initiating pore-formation in the membrane. This leads to leakage and dissipation of the proton motive force, causing lysis of the target bacterium. Although this is the general mechanism, the resulting pores can differ greatly in respect to size, stability and conductivity. Additionally, the receptor-binding and pore-formation is affected by the membrane potential and pH, implying that physiological state of the target cell also influences sensitivity. Alternative inhibition mechanisms to pore-formation have also been identified, e.g. by binding and removing lipid II from its functional site, hence inhibiting cell wall synthesis (Eijsink et al. 2002).

1.2.3 Resistance

With the early knowledge of the colicins and the notion that they attach and interact with specific receptor molecules on the surface of target cells, it was proposed that the observed resistance to colicins was due to loss of this receptor, i.e. by genetic mutations rendering the receptor unable to absorb the peptide. These receptor mutations often result in loss or poor affinity between receptor and bacteriocin. One should note that resistance is distinct from immunity, where the latter involves specific proteins protecting the producer from the activity of the bacteriocin (Hassan et al. 2012; Reeves 1965; Tagg et al. 1976). Resistance may be caused by several events, including 1) mutations altering the receptor itself, 2) mutations elsewhere affecting the placement, orientation or availability of the receptor binding site, 3) mutations causing unspecific resistance by increased afflux pump exportation or protease activity degrading bacteriocins, or 4) mutations conferring a decrease in receptor expression and thus fewer available receptor molecules (Cotter et al. 2013; Hassan et al. 2012; Lohans & Vederas 2012).

As proposed in event 2, altering the composition and structure of the cellular membrane may render a bacterium resistant to a bacteriocin by restricting the peptide from reaching its target. The physiological state of the cell can as previously mentioned also affects the effectiveness pore-formation (Eijsink et al. 2002). Additionally, lack of bacteriocin sensitivity may be a result of non-specific immunity proteins due to these proteins not always being completely specific, and thus can provide protection towards several bacteriocins. This property of cross-immunity is difficult to explain as the sequence of the immunity genes display considerable variation (Cotter et al. 2013; Eijsink et al. 2002).

1.3 Classification of bacteriocins

Bacteriocins commonly consist of 20-70 amino acids (aa) and display large variation in terms of producer species, composition, physical properties, secretion mechanism, post-translational modifications (PTMs), receptor, and inhibition spectra. Bacteria are additionally often capable of producing multiple types of bacteriocins. This complicates classification, and the issue is still under debate. New research has and will continue to alter the concept of bacteriocin classification to best represent present knowledge (Nes et al. 1996; Snyder & Worobo 2014).

Table 1.1: Classification scheme for the bacteriocins of Gram-positive bacteria. There are four main classes of bacteriocins which are further divided into various subgroups. Antimicrobial peptides produced by Gram-negative bacteria make up an entirely separate group consisting of microcins and colicins (Nes et al. 2007).

Producer bacteria	Peptide term	Class	Subclass	Description	Example
Gram-negative producers	Microcins Colicins			<10 kDa >30 kDa	
Gram-positive producers	Bacteriocins	Class I	Type A Type B Type C	Elongated and flexible Globular and rigid Multicomponent	Nisin Mersacidin Staphylococcin C55
		Class II	Subclass IIa Subclass IIb Subclass IIc Subclass IId Subclass IIe	Pediocin-like motif Two-peptide Linear, non-pediocin like Leaderless Protein-derived	Pediocin PA-1 Lactococcin G Lactococcin A LsbB Propionicin F
		Class III	IIIa IIIb	Bacteriolytic Non-bacteriolytic	>30 kDa heat-labile proteins
		Class IV		Cyclic bacteriocins	Garvicin-ML

Microcins and the larger colicins will not be addressed in detail in this thesis. These are ribosomally synthesized peptides like the bacteriocins, but are produced by Gram-negative species. They generally have a more narrow inhibition spectrum compared to the bacteriocins of the Gram-positive bacteria (Duquesne et al. 2007; Nissen-Meyer & Nes 1997). Bacteriocins produced by Gram-positive bacteria are normally not active against Gram-negative bacteria, and vice versa. The protein-sized colicins secreted by *E. coli* were for many years best studied, but the peptide-sized bacteriocins from LAB have become increasingly more interesting in recent years as good candidates for applications in various industries (Nes et al. 2007).

Class I contains the lantibiotics which are defined by their amino acids being subject to distinctive post-translational modifications (PTMs) after the peptides are ribosomally synthesized. The peptides undergo extensive PTMs such as dehydration of selected serine and threonine residues resulting in altered amino acids including 2,3-di dehydroalanine, D-alanine, 2,3-didehydrobutyrine, and formation of the characteristic lanthionine ring-structures (Eijsink et al. 2002; Hechard & Sahl 2002). Nisin was the first lantibiotic to be characterized and today more than 50 other lantibiotics have been identified. Several of these show promising activity towards MRSA and VRE. The lantibiotics are commonly divided into type A and type B lantibiotics, as well as certain two-peptide lantibiotics. Nisin is categorized as type A with its 34 amino acids and five lanthionine rings (figure 1.3). Type B lantibiotics are generally shorter and not able to facilitate pore-formation as they cannot stretch across the membrane. Despite this, they also utilize lipid II for target recognition (Martin & Breukink 2007). The lantibiotics of subgroup A are elongated, amphiphilic and positively charged, while subgroup B are globular and non-charged (Nes et al. 2007; Nissen-Meyer & Nes 1997).

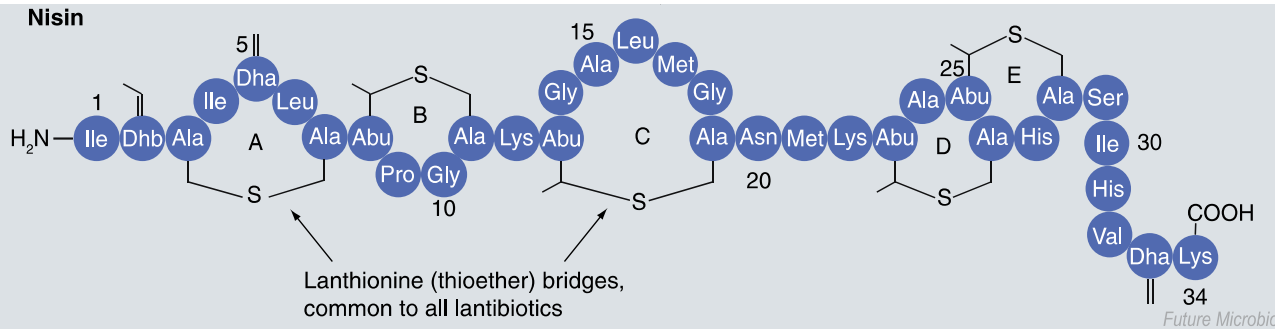


Figure 1.3: Illustration of the sequence and structure of nisin, displaying the placement of its five characteristic lanthionine rings A-E (Martin & Breukink 2007).

Class II consists of primarily non-modified peptides, and due to this they have simpler structures lacking PTMs (except for disulfide bridges or circularization, see below) as well as the genes responsible for performing distinct PTMs. Class II bacteriocins are produced by many LAB and comprise a large group, frequently divided further into five subgroups (a-e) (Nes et al. 2007).

Class IIa bacteriocins are recognizable by a pediocin-like consensus motif YGNGV and a disulfide bridge in their highly conserved hydrophilic and cationic N-terminal. The N-terminal folds into a secondary structure of antiparallel β-sheets, separated from the C-terminal by a hinge structure. The C-terminal has a secondary structure of an α-helix, and consists of a less conserved hydrophobic/amphiphilic domain involved in target specificity and membrane permeabilization (Nes et al. 2007; Nissen-Meyer & Nes 1997). The class IIa peptides consist of between 37 and 48 amino acid residues, and are known for their strong antilisterial effect. However, their activity spectra also include strains from various *Lactobacillus, Pediococcus, Enterococcus, Carnobacterium, Leuconostoc, Lactococcus and Clostridium* species (Drider et al. 2006; Eijsink et al. 2002; Fimland et al. 2005).

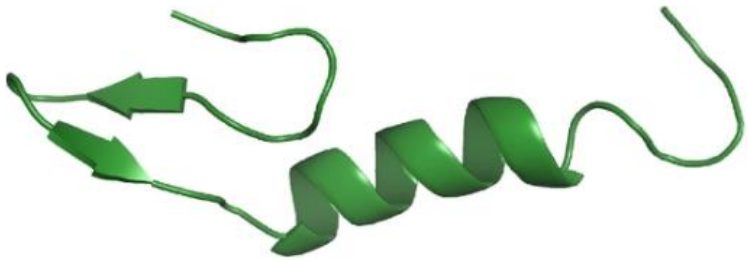


Figure 1.4: The folded structure of class IIa bacteriocin Leucocin A in solution (determined by NMR), displaying its N-terminal β-sheets and C-terminal α-helix secondary structure (Lohans & Vederas 2012).

Class IIb contains the two-peptide bacteriocins which depend on the combined action of two peptides to reach full antimicrobial activity. In some cases the individual peptides are completely inactive, while other peptides show some activity individually but combined reveals a strong synergetic effect. One-peptide bacteriocins might as well display synergetic activity when combined, but unique to the IIb bacteriocins is that their genes are located in the same operon (Eijsink et al. 2002). Additionally, the producers only express one single immunity protein, which is also located within the same operon. It has been established that the peptides attain close interaction when they inflict antibacterial activity, but the configuration of the peptides as

heterodimers is still elusive. Class IIb bacteriocins exhibit quite narrow activity spectra, and with each peptide being relatively short at 25-38 aa. Two-peptide bacteriocins also exist within the lantibiotics, but are as previously elaborated subjected to specific PTMs (Garneau et al. 2002).

Class IIc represents diverse and unsorted bacteriocins that do not fit into the other subgroups. They are non-lantibiotic, non-pediocin-like, and one-peptide bacteriocins (Eijsink et al. 2002).

Class IId are leaderless peptides, i.e. produced without leader sequence. Most bacteriocins are attached to a N-terminal leader peptide after translation, which needs to be cleaved off upon leaving the producer cell for the bacteriocin to reach functional state. The leader sequence-extension at the N-terminus is crucial to maintain inactive state whilst inside the producer, and how leaderless bacteriocin producers protect themselves from the antimicrobial activity is still unknown (Nes et al. 2007).

This study is based upon members of this subclass, and will be elaborated in more detail during subsequent sections.

Clas IIe bacteriocins are protein-derived bacteriocins produced from larger proteins by specific degradation mechanisms (Nes et al. 2007).

Class III holds large heat-labile bacteriocins with a molecular mass of > 30 kDa. Their affiliation within the bacteriocins is debated, and suggestions have been made to exclude this class and rather name these proteins as bacteriolysins (Nes et al. 2007).

Class IV is comprised of circular bacteriocins which attain a closed structure after being ribosomally synthesized as linear peptides. The linear peptide is also connected to a leader sequence that is cleaved off before formation of the mature circular structure (figure 1.5). PTMs different from those of the lantibiotics covalently link the two terminals of the cleaved peptide by inducing condensation of the N- and C-terminal amino- and carboxyl-groups, resulting in a new peptide bond. This compact structure contributes to the high stability in regard to temperature and pH observed within this class, as well as providing resistance to a number of proteolytic enzymes as result of reduced cleavage sites. Class IV bacteriocins generally exhibit broad antimicrobial activity, with members presently identified ranging from 58 to 70 aa. To this date 10 circular bacteriocins have been characterized (Gabrielsen et al. 2014).

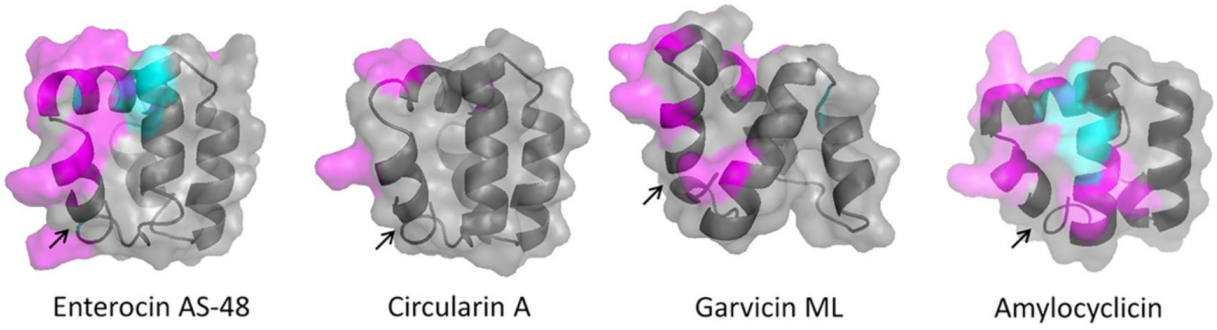


Figure 1.5: The structures of various class IV bacteriocins have been predicted using homology modeling with SWISS-MODEL, and are displayed here using PyMol. The positively charged residues are colored in magenta and negatively charged residues in cyan. The point of circularization are indicated by arrows (Gabrielsen et al. 2014).

1.4 Bacteriocins addressed in this thesis.

This particular work focuses on receptor molecules for members of the leaderless class IId bacteriocins. These are produced without leader peptides at the N-terminus, and do not undergo extensive PTMs.

1.4.1 Leaderless bacteriocin LsbB

Within the leaderless bacteriocins resides a narrow-spectrum member termed LsbB, which is produced by *L. lactis* subsp. lactis BGMN1-5 and displays specific activity exclusively towards other *L. lactis* strains. A previous study has established a zinc-dependent metallopeptidase encoded by the lactococcal *yvjB* as receptor for LsbB (Uzelac et al. 2013).

A subsequent study (Ovchinnikov et al. 2014) elucidated questions concerning the structure-function relationship of LsbB and its lactococcal target bacteria. CD spectroscopy was used to determine the optimal trifluoroethanol (TFE) and dodecylphosphocholine (DCP) conditions for obtaining maximum structuring of LsbB in solution, and NMR experiments were subsequently carried out for structural analysis of the peptide. The results concluded through superimposing the NMR structures with lowest energies that LsbB is divided into an α-helix region near the N-terminus, and a more unstructured loop region by the C-terminus (figure 1.6). The amphiphilic α-helix residues are followed by a short middle region of basic amino acids thought to facilitate the

initial and unspecific electrostatic interaction between the bacteriocin and phospholipid layer of the target cell membrane.

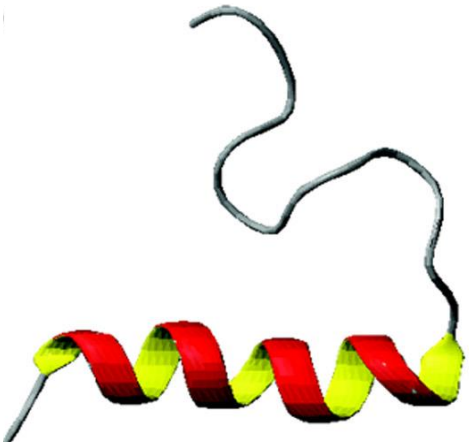


Figure 1.6: The structure of LsbB in DPC solution, represented as cartoon in Pymol from the NMR data, and showing the N-terminal region forming an alpha-helix secondary structure, with the C-terminal as a less organized loop region (Ovchinnikov et al. 2014).

Further experimental work suggested after evaluating the blocking activity of various truncated LsbB peptides that the receptor binding domain of LsbB is located in the C-terminus. Only peptides containing the C-terminus and most importantly a KxxxGxxPWE sequence motif (as from here termed the LsbB-like motif) were able to compete with LsbB by binding the receptor and inhibiting LsbB from accessing the target. In sufficiently high concentrations these truncated peptides could block the antibacterial activity completely as they do not possess antibacterial properties themselves. Additionally, alanine substitution of Trp^{25} (W25A) demolished all antimicrobial activity from LsbB, proposing that this residue is particularly crucial for successful receptor binding. These results imply that the C-terminus is responsible for initiating specific binding with the YvjB receptor, followed by disruption of the membrane integrity and pore-formation ultimately killing the target bacterium.

The consensus motif of LsbB was additionally found to be conserved across several enterococcal bacteriocins, including Ej97, EntQ and an unannotated peptide termed EntK1. Truncated C-terminus peptides derived from these enterocins could also block the activity of LsbB, which substantiates the involvement of this sequence motif in receptor binding, as well proposing that these related bacteriocins also bind to the same receptor as LsbB. Curiously, these enterocins simultaneously display a significantly broader spectrum of activity than LsbB.

The study also points out the apparent similarities between these broad-spectrum enterocins and narrow spectrum LsbB (all likely targeting the zinc-dependent metallopeptidase), and the broad-spectrum pediocin-like bacteriocins and narrow-spectrum lactococcin A (all targeting the man-PTS receptor). The extreme species specificity regarding lactococcal man-PTS exclusively functioning as receptor for lactococcin A has been established as a consequence of sequence differences between the lactococcal and non-lactococcal man-PTS genes. The same might be relevant for the lactococcal and non-lactococcal zinc-dependent metallopeptidases (Ovchinnikov et al. 2014).

1.4.2 Leaderless enterocins

As the LsbB-like motif (crucial for the activity) has been identified in several leaderless enterocins, i.e. bacteriocins with enterococcal producers, these bacteriocins are also presumed to target the same receptor. But unlike LsbB, they exhibit broad activity spectra reaching far beyond lactococcal species. The work for this thesis utilized some of these bacteriocins (with emphasis on Ej97) to investigate this conundrum.

Enterocin Ej97: 44 aa peptide produced by *Enterococcus faecalis* strains.
Enterocin K1: 37 aa peptide produced by *Enterococcus faecium* strains.
Enterocin EntQ: 34 aa peptide produced by *Enterococcus faecium* strains.

1.4.3 Sequence similarity

Despite different activity spectra and peptide lengths, all the leaderless bacteriocins mentioned above share the conserved sequence motif KxxxGxxPWE in the C-terminal, which is essential for the activity of LsbB. W25 have proved to be particularly important in this context.

```
Ej97                MLAKIKAMIKKFPNPYTLAAKLTTYEINWYKQQYGRYPWERPVA 44
K1          -------MKFKFNPTGTIVKKLTQYEIAWFKNKHGYYPWEIPRC 37
EntQ        ----------MNFLKNGIAKWMTGAELQAYKKKYGCLPWEKISC 34
LsbB        --------------MKTILRFVAGYDIASHKKKTGGYPWERGKA 30
                                :   ::  ::  .*:: *  ***   .
```

Figure 1.7: Multiple sequence aliment (MSA) of the sequences of Ej97, K1, EntQ and LsbB, predicted by ClustalW. Fully conserved amino acid residues are marked with "*" and less conserved residues with ":" or "." depending on the degree of amino acid similarity. The conserved LsbB-like motif is marked with red.

1.5 Bacteriocin receptors and strategies used in their identification

Due to the rapid emergence of antibiotic resistant bacteria, it is desirable to identify antimicrobial agents that target different sites than the presently utilized antibiotics in order to avoid problems with cross-resistance. The bacteriocins, targeting specific receptors in the cell membrane and exerting their activity by pore-formation are interesting in this aspect. Knowledge on their receptors for target recognition is thus desirable for further development in the field and attracting the attention of various industries for commercial distribution. Understanding receptor-interaction is also important for engineering of improved bacteriocins and hybrids (Cotter 2014). As will be described in this section, various strategies have been applied to elucidate receptor molecules for bacteriocins over the past decades.

The class I lantibiotic nisin and Lipid II

Nisin was the first lantibiotic to be identified, and is produced by *L. lactis* strains. In 1998, a precursor of peptidoglycan in the bacterial cell wall, lipid II, was identified as the docking molecule for nisin. It had already been assumed that nisin disrupts the cellular membrane of target bacteria due to exposure to nisin inducing rapid efflux of cytoplasmic components from target cells. Nisin specifically belongs to a subgroup of lantibiotics which are recognized by an elongated and cationic structure consisting of five lanthionine rings A-E. By studying the activity of mutated nisin molecules, it was discovered that the two first lanthionine rings A and B at the N-terminal are essential for binding lipid II, while the whole molecule of five rings is necessary to form lethal pores in the membrane by using the C-terminal domain. It was also observed that nisin has two mechanisms for killing target cells, either by pore formation, or by blocking lipid II from being incorporated into the peptidoglycan layer and thus inhibiting cell-wall synthesis (Hechard & Sahl 2002; Wiedemann et al. 2001). Lipid II is a cell-wall precursor of the peptidoglycan layer found in both Gram-positive and Gram-negative bacteria. Gram-positive cells acquire a thick layer of peptidoglycan, while Gram-negative cells have a thinner layer and an additional outer membrane. Nisin exhibits a broad spectrum of activity, which is reasonable as lipid II is such a common component. Several other lantibiotics such as subtilin and the epidermin family also share this conserved ring system, suggesting that they also bind and employ lipid II for target interaction. The epidermins do however not have the sufficient peptide length to span bacterial membranes, and pore formation is thereby not possible. Experiments

have suggested that these peptides instead exert their activity by binding lipid II and inhibiting peptidoglycan synthesis by removing lipid II from its active site. This removal of lipid II from its functional site is possibly the main mechanism of lantibiotics which are too short to form stable pores (Martin & Breukink 2007).

The class IIc lactococcin A and the class IIa pediocin-like bacteriocins, and the man-PTS

Lactococcin A (class IIc) of *L. lactis* is exclusively active against other *L. lactis* species. Past studies have revealed that lactococcin A utilizes a permeabilizing mechanism involving the mannose phosphotransferase system (man-PTS) for target cell recognition and interaction. The man-PTS is responsible for coupling carbohydrate import with phosphorylation of the sugar molecules inside bacterial cells, and is not present in eukaryotes. The man-PTS consists of the enzymes EI, HPr and the carbohydrate specific EII. EII is composed of IIAB, IIC and IID, where the two latter are located in the membrane and involved in recognition by bacteriocins. The IIAB component located in the cytoplasm is on the other hand not involved in sensitivity. The man-PTS genes in different bacteria exhibit subtle sequence variations that makes it possible for the bacteriocins to differentiate species (Kjos et al. 2009). Lactococcin A producers maintain self-immunity by expression of a specific immunity protein LciA which form a complex with the bacteriocin and receptor components to prevent the host from being killed. This complex appears to occur only in the presence of the bacteriocin. By using tagged immunity-proteins (fLciA) and the cognate bacteriocin, it was possible to co-purify a protein complex consisting of immunity protein, receptor proteins, and bacteriocin from cell extracts by immunoprecipitation. From this complex the man-PTS components were identified with mass spectrometry (MS) and peptide mass fingerprinting. Additionally, lactococcin A resistance was established in strains where the genes encoding the IIC and IID transmembrane proteins of the man-PTS had been deleted. The study finally established that heterologous expression of the lactococcal man-PTS operon in the naturally resistant *Lactobacillus sakei* rendered this species susceptible to lactococcin A. To identify exactly which components of the man-PTS are used for target recognition by lactococcin A, the genes were cloned and expressed as single genes or pair of genes using the NICE system (nisin-controlled gene expression) to regulate expression. Only expression of the gene pair ptnC (for IIC) and ptnD (for IID) conferred bacteriocin sensitivity (Diep et al. 2007). Various other class II bacteriocins including lactococcin B (class IIc) and pediocin-like bacteriocins (class IIa)

also exploit the IIC and IID components of the man-PTS in killing sensitive target cells. In the case of pediocin-like class IIa bacteriocins, it has been further discovered that these bacteriocins specifically interact with a ~20 aa extracellular loop in the IIC component to bind and initiate pore-formation (Diep et al. 2007; Kjos et al. 2009; Kjos et al. 2010).

The circular class IV bacteriocin garvicin ML and the maltose ABC transporter

More recent developments in the field include the discovery of the maltose ABC transporter as receptor for the circular class IV bacteriocin garvicin ML. This bacteriocin is produced by *Lactococcus garvieae* DCC43, an isolate from the intestine of Mallard ducks (Borrero et al. 2011). Spontaneous *L. lactis* 1403 mutants with lowered sensitivity to garvicin ML were generated and their fermentation profile was assessed for discovering possible disruption in sugar metabolism. This resulted in observations that the mutants were unable to grow effectively on starch and maltose sugar, suggesting that the mutants had lost their ability to metabolize these carbohydrates. This phenotype seemed to somehow be linked to the garvicin ML resistance and thus its receptor. Genotype analysis by sequencing revealed that the mutants had a chromosomal deletion of 13.5 kb spanning 12 open reading frames (ORFs). The data showed that this section included the genes *malEFG*, encoding a membrane-bound maltose ABC-transporter. Additional growth studies showed that the sensitivity to Garvicin ML was elevated when wild type cells were grown on maltose, in contrast to when the bacteria had other available resources such as glucose (Gabrielsen et al. 2012).

The two-peptide class IIb bacteriocin lactococcin G and the UppP receptor

The most recent discovery in target receptor recognition was made on the two-peptide bacteriocin lactococcin G produced by certain *L. lactis* strains. By employing whole genome sequencing (WGS), mutations within the *uppP* gene were located in 12 resistant *L. lactis* mutants. This gene encodes a membrane spanning undecaprenyl pyrophosphate phosphatase, which functions as a component involved in peptidoglycan synthesis. Using bioinformatics tools for SNP detection, the resistant phenotype of these mutants showed correlation to mutations within or up-stream *uppP*. The mutants were affected by a variety of serious mutations, including premature stop codons, mutations in regulatory regions, and amino acid substitutions. Lactococcin G is specifically potent to *L. lactis* strains and requires the cooperation of two

separate peptides (a 39 aa α-peptide ad a 35 residue β-peptide) to exert antimicrobial activity. The importance of this gene in sensitivity was confirmed by heterologous expression of lactococcal *uppP* in *Streptococcus pneumonia*, a naturally resistant strain as lactococcin G can only target the lactococcal version of this gene. For unknown reasons, the researchers were unable to reintroduce the intact *uppP* gene to resistant *L. lactis* mutants despite several trials using various cloning vectors. Instead, the approach was modified to revolve around heterologous expression of the lactococcal wild type *uppP* in *Streptococcus pneumonia*. This induced sensitivity to lactococcin G in the naturally resistant host. This represents the first time a receptor have been identified for a two-peptide bacteriocin. The study also revealed that another class IIb bacteriocin, enterocin 1071, also exploits UppP for target recognition. The work also included a fermentation assay to study potential differences in sugar metabolism between wild type and mutants, but did not give any positive results, implying that the receptor was not involved in sugar metabolism pathways. One mutant also had an additional mutation in *oppD*, the ATPase component of an ABC-type transporter system (Cotter 2014; Kjos et al. 2014).

This WGS-based receptor identification strategy (figure 1.8) proves to be a good alternative as the cost of sequencing is on the decrease, and as the heterogeneity of bacteriocins substantiates a large variety of different receptor genes remains to be identified (Cotter 2014).

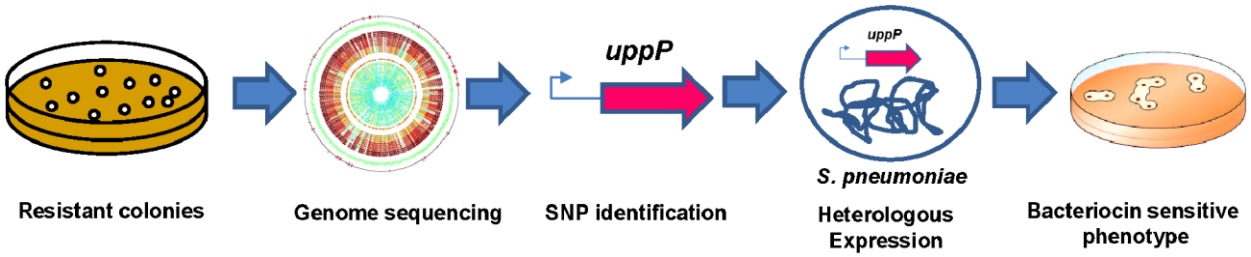


Figure 1.8: Visualization of the pipeline utilized in the identification of UppP as receptor for lactococcin G. WGS was the first step after generation of resistant mutants, and by modern SNP identification tools a candidate receptor gene was established using a bioinformatics approach. The involvement of the candidate gene was subsequently confirmed by experimental evidence through heterologous expression of the wild type gene, which introduced sensitivity to lactococcin G in a naturally resistant host (Cotter 2014).

The leaderless class IId bacteriocin LsbB and a Zn-dependent metallopeptidase

Another recent finding in bacteriocin receptors came with the identification of a Zn-dependent metallopeptidase coded by the *yvjB*-gene in *L. lactis* as receptor for the leaderless class IId

bacteriocin LsbB. YvjB is a lactococcal member of the RIP metallopeptidase RseP protein family and the M50 superfamily. The strategy employed was to create a cosmid library of an LsbB-sensitive strain and experiment with which fragment from this library restored sensitivity to the bacteriocin when expressed in resistant mutants. Each cosmid library contained a fragment and the lambda phage cos-sequence to make it suitable as a cloning vector. The results showed that a cosmid carrying a 40 kb insert was able to restore sensitivity, and further subcloning of this cosmid to narrow the candidate gene possibilities revealed an open reading frame on a 1.9 kb fragment as a potential receptor gene. This open reading frame encodes the zinc-dependent protease enzyme YvjB. The involvement of this particular gene as receptor was confirmed by additional whole genome sequencing showing *yvjB* mutations in all mutants, deletion of functional *yvjB* inducing resistant phenotype, and homologous expression of wild type *yvjB* in naturally resistant hosts introducing sensitivity to LsbB (Uzelac et al, 2013).

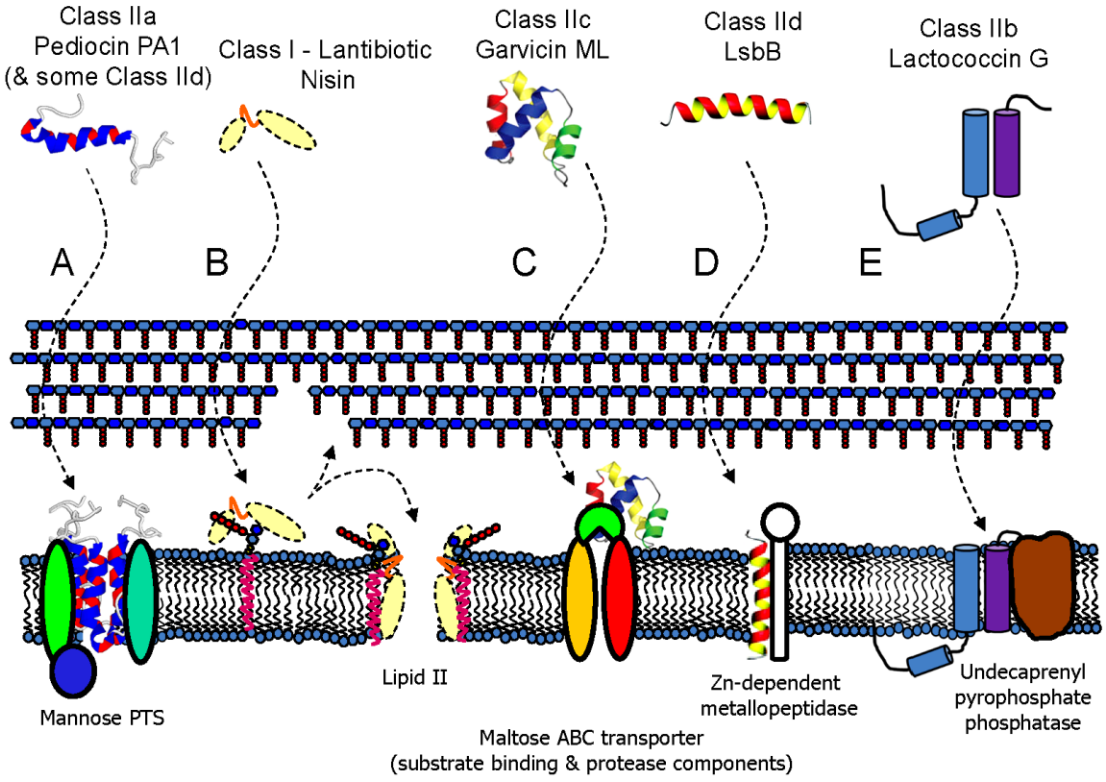


Figure 1.9: A depiction of the presently identified bacteriocin receptors initiating pore-formation in susceptible target cells. Class IIa and some IIc bacteriocins exploits components of the man-PTS, nisin and other lantibiotics target lipid II, circular class IV garvicin ML attacks the Maltose ABC transporter, class IId leaderless bacteriocin LsbB bind a zinc-dependent metallopeptidase, and finally recent developments have established UppP as docking molecule for the two-peptide class IIb bacteriocin lactococcin G (Cotter 2014).

1.6 The zinc-dependent metallopeptidase Rsep of the M50-superfamily

As previously elaborated, lactococcal RseP (a zinc-dependent metallopeptidase YvjB) of the M50 superfamily has been identified as the receptor for the leaderless bacteriocin LsbB. Members of this family exert proteolytic activity, and such protease enzymes serve either as cellular housekeepers by removing damaged proteins in the cell, or as part of a regulatory pathway by cleaving specific regulatory components. The latter are additionally termed RIP (regulated intramembrane proteolysis) enzymes. RseP is such a RIP enzyme.

In studies on *E. coli*, the *rseP* gene is well established as involved in initiating stress responses by cleaving an anti-sigma factor and thereby allowing cognate sigma factors to induce transcription of stress response genes (figure 1.10). RseP is in this manner implemented in regulatory proteolysis by a proteolytic cascade reaction which is initiated in response to extracellular stress signals. By default, the sigma factor σ^E is bound to its regulatory anti-sigma factor RseA. Stress signals lead to the RseP cleaving the anti-sigma factor RseA in accordance with cleavage by the inner membrane protease DegS. These cleavages release the cytoplasmic domain of RseA into the cell where another protease ClpXP degrades the remaining parts of the RseA, and thus frees σ^E. This enables σ^E to bind DNA and RNA polymerase in the promotor site upstream the genes responsible for stress response. The process is regulated in this manner to inhibit unnecessary production of stress response genes when there is no stress present cellular environment. The *rseP* has orthologous genes in various species (Akiyama et al. 2004; Kroos & Akiyama 2013).

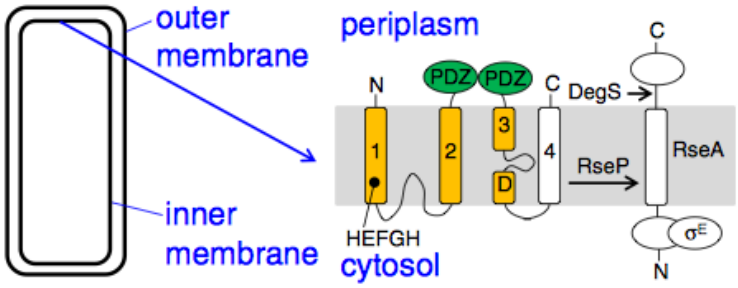


Figure 1.10: DegS cleaves the periplasmic C-terminal domain of anti-sigma factor RseA, before RseP cleaves the trans membrane region (displayed by arrows), which release the N-terminal domain of RseA and the connected σ^E into the cell, where the remaining RseA is finally degraded by ClpXP in the cytoplasm. Thus, releasing σ^E to initiate transcription of stress response genes (Kroos & Akiyama 2013).

1.7 Aim of this study

The main goal of this thesis was to identify the target receptor for the leaderless bacteriocin Ej97. The intriguing fact that sequence characteristics of Ej97 substantiates that it likely exploits the LsbB-receptor, while simultaneously possessing the ability to dock with many other bacterial species LsbB itself is unable to attack, made this an interesting task.

The work completed in this study centered on the following tasks:

1. A spot-on-lawn inhibition spectrum assay for evaluating the activity of three homologous bacteriocins LsbB, K1 and Ej97 towards approximately 50 indicator bacteria. Indicators with naturally resistant colonies occurring within the inhibition zones would be the basis of further work progress.
2. Accumulation of resistant mutants and utilization of microtiter plate assays with two-fold bacteriocin dilutions to assess the resistance level (phenotype) of these mutants relative to the wild type.
3. Sequencing of the putative receptor RseP, recognized as a membrane-associated metalloprotease and homologous with LsbB receptor YvjB in *L. lactis* IL1403, in order to identify mutations (genotype) causing the observed resistant phenotypes.
4. Whole genome sequencing for uncovering reasons for the observed resistance in isolates lacking mutations within *rseP*, and assessment of possible unspecific resistance mechanisms present in these resistant mutants.

2. MATERIALS

2.1 Growth media and agars

Medium	Supplier
BHI (Brain-Heart-Infusion)	Oxoid

Growth medium: 9.25 g BHI

dH_2O to 250 ml

Sterilized in autoclave for 15 min at 121 °C

Soft agar: BHI medium with agar (8 g/L)

Agar: BHI medium with agar (15 g/L)

MRS (de Man, Rogosa, Sharpe) Oxoid

Growth medium: 13 g MRS

dH_2O to 250 ml

Sterilized in autoclave for 15 min at 121 °C

Soft agar: MRS medium with agar (8 g/L)

Agar: MRS medium with agar (15 g/L)

2.2 Bacterial strains

Table 2.1: Bacterial strains implemented in the work for this thesis.

Bacterium	**LMG lab strain reference code**
Bacillus cereus	LMG 2805
Enterococcus avium 208	LMG 3465
Enterococcus faecalis UI50	LMG 2333
E. faecalis 2708 RA	LMGT 3358*
E. faecalis DEC23	LMGT 3386
E. faecalis SMF37	LMGT 3370

Enterococcus faecium L50	LMG 2763
E. faecium P21	LMG 2783
E. faecium AL41	LMG 2876
E. faecium	LMG 2787*
Enterobacter sp 366	LMG 3287
Escherichia coli ATCC 14763	LMG 2746
E. coli JM 109	LMG 3235
Lactococcus strain F4-13	LMG 2070
Lactobacillus curvatus 89	LMG 2355
L. curvatus CTC 435	LMG 2371
L. curvatus	LMG 2353
L. curvatus NCFB 2739 B	LMG 2715
Lactobacillus plantarum ssp 965	LMG 2003
L. plantarum R	LMG 2352
L. plantarum NC-8	LMG 2357
L. plantarum C11	LMG 2358
L. plantarum C11	LMG 2362
L. plantarum	LMG 2379
L. plantarum	LMG 3125
Lactobacillus sakei 706	LMG 2334
L. sakei LS3	LMG 2356
L. sakei 148	LMG 2361
L. sakei	LMG 2380
L. sakei 791	LMG 2799
Lactococcus garvieae 1546	KS1556*
L. garvieae DCC43	LMG 3390*
Lactococcus lactis IL1403	LMG 2705
L. lactis F14	LMG 2081
L. lactis 1403	LMG 2130**
L. lactis QU5	LMG 3419
Leuconostoc gelidium Ta	LMG 2386
Listeria innocua	LMG 2710
L. innocua	LMG 2785
Listeria ivanovii Li4	LMG 2813
Listeria monocytogenes EGDe	LMG 2604
L. monocytogenes 279 serotype 4	LMG 2650
L. monocytogenes 400 serotype 4	LMG 2651
L. monocytogenes	LMG 2652
L. monocytogenes 223 serotype 1	LMG 2653
Pediococcus acidilactici Pac1.0	LMG 2002*
Pediococcus pentosacens FBB61.1	LMG 2001
Pediococcus pentosaceus PPE 1-2 M	LMG 2366
P. pentosaceus FBB 63 B	LMG 2722
Staphylococcus aureus	LMG 3022
S. aureus	LMG 3023
S. aureus	LMG 3242
Streptococcus salivarius	LMG 1301

* Strains where bacteriocin resistant mutants arose. Further experimental work revolved around these.

** Strain used both as indicator in inhibition assay and for production of control bacteriocin nisin.

2.3 Synthetic peptides

Table 2.2: Bacteriocins used in experiments. Ordered from GenScript.

Bacteriocin	Amino acid sequence	Reference
Ej97	MLAKIKAMIKKFPNPYTLAAKLTTYEINWYKQQYGRYPWERPVA	Galvez et al. 1998
K1	MKFKFNPTGTIVKKLTQYEIAWFKNKHGYYPWEIPRC	Ovchinnikov et al. 2014
EntQ	MNFLKNGIAKWMTGAELQAYKKKYGCLPWEKISC	Cintas et al. 2000
LsbB	MKTILRFVAGYDIASHKKKTGGYPWERGKA	Gajic et al. 2003
BHT-B	MWGRILAFVAKYGTKAVQWAWKNKWFLLSLGEAVFDYIRSIWGG	Hyink et al. 2005
Gar-ML	MFDLVATGMAAGVAKTIVNAVSAGMDIATALSLFSGAFTAAGGIMA LIKKYAQKKLWKQLIAA	Borrero et al. 2011

2.4 Laboratory equipment

Equipment	Supplier
96-well PCR plates	VWR
Acid-washed glass beads (<106 microns)	Sigma
Cryo-tubes 2 ml	Sarstedt
Culture tubes 10 ml	-
Eppendorf tubes 2 ml	Eppendorf
FastPrep tubes	-
Gel-electrophoresis equipment (rack, molding form, comb)	Bio-Rad
Gel photo system with UV spectrum	UVP
Glass bottles 500 ml for autoclaving	-
Gloves	VWR
Magnetic stand	-
Microseal for 96-well microtiter plates	-
Microtiter 96-well plates	Sarstedt
Multi channel pipettes	Thermo Scientific
Parafilm	Bemis
PCR tubes 0.2 ml	-
Petri dishes	-

Pipettes	Eppendorf
Pipette tips	VWR
Plastic loops/streakers	Sarstedt
Qubit 0.5 ml tubes	Life Technologies

2.5 Instruments

Instrument	Supplier
Autoclave	Matachana
Digital weight	Salter
Electrophoresis electricity supply	Hoefer Scientific Instruments
Eppendorf centrifuge	Eppendorf
FastPrep machine	Savant
Freezer (-80°C)	Forma Scientific
Gas burner	Intergra Biosciences
Heating block for 1.5 ml centrifuge tubes	Stuart Scientific
Illumina MiSeq	Illumnia
Incubator (30 °C)	Fermaks
Microplate centrifuge	Eppendorf
Microtiter plate scanner	Illumina
Microtiter plate shaker	InterMed
NanoDrop ND-1000	Nanodrop Technologies
PCR machine	Bio-Rad
Qubit 2.0	Life Technologies
Sterile bench w/fume hood	Holten Laminaire
Vortex	Scientific Industries
Water bath	Julabo

2.6 Software

Software	Supplier
BioEdit	Tom Hall

CLC workbench 7.5 and 5.5	Qiagen
Geneious 7.1.5	Biomatters Ltd.
Illumina BaseSpace	Illumina
Mauve aligner 2.3.1	Genome Evolution Lab.
NanoDrop 3.0.0	Thermo Scientific
PROKKA annotation	VBC
RAST annotation	NMPDR

2.7 Kits

The protocols accompanying the various kits were utilized in their designated procedures, but usually with certain alterations. Thus all procedures in the methods section will be described in in their entirety.

Kit	Supplier
Agencourt AMPure XP 60 ml kit (A63881)	Beckman Coulter
AMPure XP – 60 ml Speed Beads – Mag Carboxyl – 1000 ml	
E.N.Z.A™ Plasmid MiniPrep Kit (D6942-02)	Omega
Solution I	
Solution II	
Solution III	
Buffer HB	
DNA wash buffer	
Elution buffer	
RNase A	
HiBind DNA mini columns	
2 ml collection tubes	
GenElute^{TM} Bacterial Genomic DNA Kit (NA2120)	Sigma-Aldrich
Gram positive Lysis Solution	
Lysis Solution T	
Lysis Solution C	

Wash Solution 1	
Wash Solution Concentrate	
Elution Solution (10 mM Tris-HCl, 0.5 mM EDTA, pH 9.0)	
Column Preparation Solution	
Proteinase K	
RNase A Solution	
GenElute Nucleic Acid Binding Columns in tube (red)	
2 ml collection tubes	
MiSeq Run Reagents (V3-600 cycles) (MS-102-3003)	Illumina
Hybridization Buffer HT1	
Incorporation Buffer	
PR2 Bottle	
MiSeq Flow Cell (single use)	
Reagent Cartridge (single use)	
Nextera XT DNA Sample Preparation Kit (FC-131-1096)	Illumina
Amplicon Tagment Mix, 24 RXN (ATM)	
Tagment DNA Buffer (TD)	
Nextera PCR Master Mix (NPM)	
Resuspension Buffer (RSB)	
Library Normalization Additives 1 (LNA1)	
Library Normalization Wash 1 (LNW1)	
Hybridization Buffer (HT1)	
Neutralize Tagment Buffer (NT)	
Library Normalization Beads 1 (LNB1)	
Library Normalization Storage Buffer 1 (LNS1)	
Nextera XT Indexing Kit (96-indeces) (FC-131-1002)	Illumina
Index Primers (S501-S508)	
Index Primers (N701-N712)	
Nucleospin® PCR Clean-up Gel Extraction Kit (740609.250)	Macherey-Nagel
Binding Buffer NTI	
Wash Buffer NT3	

Elution Buffer	
NucleoSpin® Gel and PCR Clean-up Columns	
2 ml collection tubes	
PhiX control V3 kit (FC-110-3001)	Illumina
PhiX control	
TruSeq Index Plate Fixture Kit (FC-130-1005)	Illumina
Plate Fixture	
Qubit™ dsDNA HS (high sensitivity) Kit (Q32854)	Invitrogen
dsDNA HS dye reagent concentrate	
dsDNA HS Buffer	
dsDNA HS Standard 1	
dsDNA HS Standard 2	

2.8 Chemicals and reagents

Chemical	Supplier
10x Taq Buffer (-MgCl2)	Life technologies
50xTAE*	-
Agar	Merck
Agarose	Life technologies
dH_2O	Produced locally
Ethanol	Arcus
GelRed	Biotium/VWR
Glycerol	Merck
Loading buffer 6x	New England BioLabs
$MgCl_2$ 50 mM	Life technologies
Tween 20	Sigma

* 50xTAE (Tris-acetate-EDTA): 242 g Tris base is added to 57.1 ml acetic acid and 100 ml 0.5 M EDTA pH 8.0. dH_2O is added to a final volume of 1 liter.

2.9 Enzymes

Protein/enzyme	Supplier
Taq® DNA polymerase	Life technologies
Proteinase K 20 µg/ml	Sigma
RNase 10 mg/ml	Sigma
Lysozyme L4919 from chicken egg white	Sigma

2.10 DNA and standards

DNA	Supplier
1kb ladder (50 ng/µl)	New England BioLabs
Deoxynucleotides (dNTPs 10 mM)	Life technologies

2.11 Primers

All primers were ordered from Life Technologies

Table 2.3: Primer sequences designed specifically for this thesis.

Primer	Sequence (5'→3')	Application
Ped_F	TGAATCTGCCCTTAAGCGGTTC	Forward primer for amplification and sequencing of *rseP* in *P. acidilactici* LMG 2002
Ped_M	GACGTACAGTTCCAATCAGCATC	Middle primer for sequencing of *rseP* in *P. acidilactici* LMG 2002
Ped_R	TGTGACTAAGCACTTCCGCATC	Reverse primer for amplification and sequencing of *rseP* in *P. acidilactici* LMG 2002
Ent_F	CGAAGTGGTCAAGTCCAATGGT	Forward primer for amplification and sequencing of *rseP* in *E. faecalis* LMG 3358
Ent_M	GTGCGGATTGCGCCACTTGAC	Middle primer for sequencing of *rseP* in *E. faecalis* LMG 3358
Ent_R	GATGACTTAAGACTTCTGCATCAT	Reverse primer for amplification and sequencing of *rseP* in *E. faecalis* LMG 3358
DCC_F	ACCTGGACATGGAGGCGTACT	Forward primer for amplification and sequencing of *rseP* in *L. garvieae* DCC43

DCC_M	CAGTTGACCATGATGCAACGATC	Middle primer for sequencing of *rseP* in *L. garvieae* DCC43
DCC_R	CGCACATAGCCTGCACGTACT	Reverse primer for amplification and sequencing of *rseP* in *L. garvieae* DCC43
Gar_F	TGCAGGAGCCAAGAGTTCTACT	Forward primer for amplification and sequencing of *rseP* in *L. garvieae* 1546
Gar_M	TGAGTGATACCAATGCGGTATCT	Middle primer for sequencing of *rseP* in *L. garvieae* 1546
Gar_R	GGCTTGTTCTAAGCCTTGAGGT	Reverse primer for amplification and sequencing of *rseP* in *L. garvieae* 1546
EF_F	GCTCTTAGCAAGATTTGATGGC	Forward primer for amplification and sequencing of *rseP* in *E. faecium* LMG 2787
EF_M	CGTCCACACTGACTACCTCATC	Middle primer for sequencing of *rseP* in *E. faecium* LMG 2787
EF_R	CTTAGACCGTTTCGACAGTTTGC	Reverse primer for amplification and sequencing of *rseP* in *E. faecium* LMG 2787
EF_R2	TGCAATCTGTCGACGTGACAC	Additional reverse primer for sequencing of *rseP* in E. faecium LMG 2787-mutants with transposons.
T1_F	AGCTAGCTCAAAGGAAGAGGC	Additional forward primer for sequencing *E. faecium* LMG 2787-mutants B32 and B80 with transposon type I.
T2_F	TGCAATCTGTCGACGTGACAC	Additional forward primer for sequencing *E. faecium* LMG 2787-mutants B74, B76 and B84 with transposons type II.
T3_F	GCTCGAACAGCTAAGAATGCCT	Additional forward primer for sequencing *E. faecium* LMG 2787-mutants B30, B70 and B94 with transposon type III
3358wt	CGATTAATAATCAAAAAATCAAAAAATAC	Mutation frequency analysis using repeat specific primers for *E. faecalis* LMG 3358 wild type.
Ins_62	TCAAAAAATCAAAAAATCAAAAAATAC	Mutation frequency analysis using repeat specific primers with a melting temp. of approximately 62 °C
Ins_68	TAATCAAAAAATCAAAAAATCAAAAAATAC	Mutation frequency analysis using repeat specific primers with a melting temp. of approximately 68 °C
Del_62	CTTATCGATTAATAATCAAAAAATAC	Mutation frequency analysis using repeat specific primers with a melting temp. of approximately 62 °C
Del_68	GTCTTATCGATTAATAATCAAAAAATAC	Mutation frequency analysis using repeat specific primers with a melting temp. of approximately 68 °C
Ent_R2	GTACTTACTCCAGCATTGGATGC	Reverse primer used with the mutation specific forward primers to yield amplicons of ~350 bp

3. METHODS

3.1 General methods in microbiology

3.1.1 Sterile working technique

Throughout the work implemented in this thesis, procedures with high risk of contamination were consistently carried out in sterile work benches accompanied by diligent use of gloves, ethanol and gas burners to avoid compromising the bacterial samples. Tools and equipment were autoclaved and/or sterilized with ethanol and gas burner prior to contact with samples.

3.1.2 Preparation of growth medium and agars

Cultivation and experimentation with bacteria require suitable growth media, soft agars and agars. These were assembled according to recipes provided by the manufacturer, followed by autoclaving and appropriate storage for later use. Growth media for cultivation of pure cultures were transferred in 5 ml dosages to culture tubes prior to autoclaving in order to minimize the risk of contamination. Agars were cooled to 50-60 °C and transferred to petri dishes before being used for streaking bacteria to single colonies, or preforming spot-on-lawn experiments. Soft agars in liquid state were employed to dilute and spread bacteria evenly onto solid agar.

3.1.3 Streaking bacteria onto agar plates

To achieve single colonies which were later used for making pure cultures, bacteria from the LMG glycerol stock were streaked onto agar plates containing solidified growth medium. This was executed by using sterile plastic loops and a streaking pattern aiming to dilute the bacteria to single colonies. The plates were incubated at 30 °C overnight or for the amount of time required for substantial bacterial growth to appear. Plates were subsequently stored at 4 °C for further use.

3.1.4 Inoculation and cultivation of overnight pure cultures (ON-cultures)

Fresh pure cultures were attained by inoculating single colonies from agar plates using sterile plastic loops or toothpicks, and transferred to 5 ml of befitting growth medium in autoclaved culture tubes. After incubation overnight at 30 °C, the pure ON-cultures were homogenized on vortexers before being applied in experimental work.

3.1.5 Storage of bacterial isolates (glycerol stock)

For long-term storage of bacterial isolates, stock samples were assembled in 2 ml cryo-tubes by mixing 0.5 ml ~ 50 % glycerol solution (diluted from 85% glycerol with dH_2O) with 1.0 ml pure ON-culture to obtain an concentration of 15-20 % glycerol in the stock samples. Glycerol is necessary to prevent formation of ice crystals resulting in cell disruption at low temperatures. The components were merged thoroughly by pipetting and stored at -80 °C.

3.2 Schematic of work progression

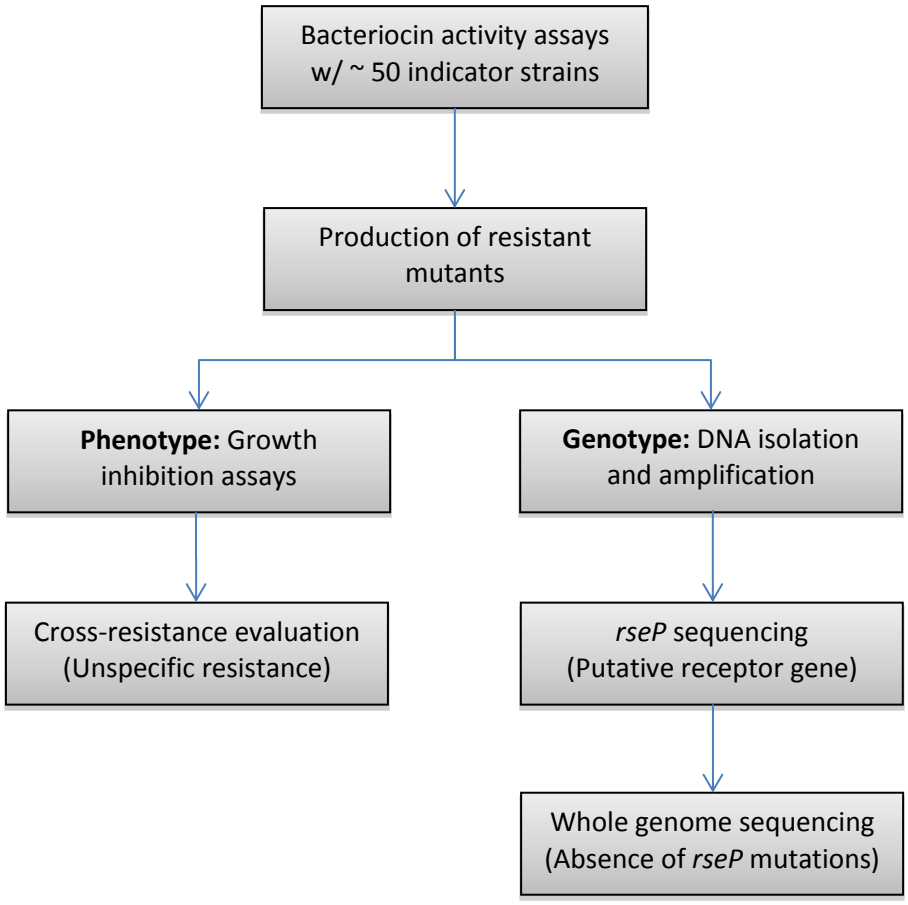


Figure 3.1: Workflow overview of the proceedings for this thesis. The initial production of Ej97-resistant mutants was followed by various assays aiming to pinpoint genotype variations explaining the observed phenotype variations.

3.3 Spot-on-lawn inhibition spectrum assay

To assess the activity spectra of LsbB, Ej97 and K1, an inhibition spectrum assay was conducted employing approximately 50 indicator bacteria from various genera (table 2.1). The strains had in advance been inoculated and grown to pure cultures overnight at 30 °C, commonly utilizing BHI medium for this purpose. MRS medium was applied for optimal and selective growth conditions (contains sodium acetate) when cultivating lactobacilli strains. Work with pathogenic strains in the inhibition spectrum assays was executed separately in a designated pathogen lab.

In preparation of the inhibition spectrum assay, soft agar was re-heated using a microwave and kept on water bath at ~ 55 °C to sustain liquid state throughout the procedure, while bacteriocins and proteinase K were kept on ice to prevent degradation. The bottom of the agar plates were marked according to indicator strain and the intended placement of bacteriocins and proteinase K as shown in figure 3.1. Proteinase K is an enzyme which digests proteins, and was implemented in the experiment to confirm that the inhibitory effects observed was due to the activity of a protein-like structure and not caused by pH-decrease or other inhibitory effects. Proteinase K will break down bacteriocins and thus inhibit killing of the indicator strain in the area it is applied to, causing disruption of the otherwise circular inhibition zone created by the bacteriocin.

In addition to the synthetic bacteriocins LsbB, Ej97 and K1, ribosomally produced nisin by *L. lactis* 1403 LMG 2130 ON-culture was added as a control for an LsbB non-related bacteriocin in the experiment due to possessing a broad and well-known activity spectrum.

<u>Materials:</u>
5 ml indicator strain ON-culture (table 2.1)
5 ml nisin producer *L. lactis* 1403 ON-culture of (table 2.1)
BHI agar plates
BHI soft agar
Ej97 0.5 µg/ml
K1 0.5 µg/ml
LsbB 0.5 µg/ml
Proteinase K

Procedure:

1. 100 µl of indicator ON-culture was thoroughly merged into 5 ml of soft agar in a sterile culture tube using vortexer to attain a 50x dilution. The solution was immediately and carefully poured onto the correspondingly marked agar plate and evenly distributed on the surface before being allowed to solidify for approximately 10 minutes.
2. 4 µl of each of the bacteriocins EntEj97, EntK1 and LsbB, as well as the nisin producer, were then added to the indicated locations on the plate.
3. 2 µl proteinase K was subsequently applied to the marked placements in close proximity of the bacteriocin droplets.
4. Finally, the plates were once more allowed to dry before being incubated at 30 °C overnight for bacterial growth and inhibition zones to appear.

The resulting growth patterns were photographed on the following day and analyzed by evaluating the presence and relative sizes of the inhibition zones induced by the bacteriocins. Apparent resistant mutants appearing within the inhibition zones were to be the foundation of further work for the project.

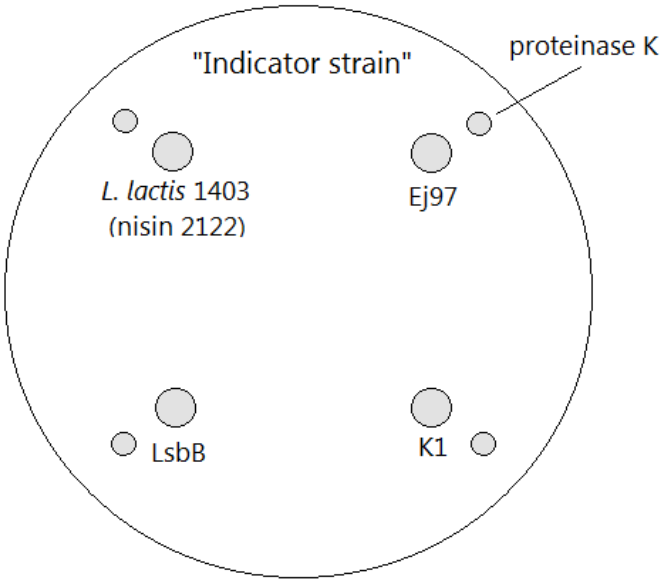


Figure 3.1: Reagent placements for the inhibition spectrum assay, illustrated as an agar plate viewed directly from above. Proteinase K was added to the smaller indicated spots next to each of the bacteriocins to confirm that the eventual inhibitory activity observed on the indicators was due to substances of proteinaceous nature, i.e. bacteriocins.

3.4 Accumulation of bacteriocin resistant mutants

Indicator bacteria where resistant colonies had appeared during the inhibition spectrum assay were applied in new spot-on-lawn experiments using varying bacteriocin concentrations (0.1 µg/ml, 0.25 µg/ml and 0.5 µg/ml), aiming to produce more mutants. After overnight incubation at 30 °C, resistant colonies were picked from the plates using sterile toothpicks and transferred onto fresh agar plates before being incubated at 30 °C overnight. Finally, plastic loops were used to once more streak the resistant bacteria onto fresh agar plates to achieve single colonies. These would be used to cultivate pure cultures for microtiter plate assays and DNA extraction. The plates were stored at 4 °C and the pure cultures as glycerol stock at -80 °C.

Additional mutants resistant to Ej97 or K1 and originating from *L. garvieae* DCC43 and *L. garvieae* 1546 were supplied by Kirill Ovchinnikov, as well as *E. faecium* 2787-mutants resistant to K1 or EntQ from Paweł Oskólski.

3.5 Microtiter plate assays with two-fold bacteriocin dilution

To estimate the degree of resistance in the isolated mutants compared to the wild type, growth experiments were carried out using microtiter plates with two-fold bacteriocin dilutions. This allowed for calculations of minimum inhibitory concentration for reducing growth by > 50% (MIC_{50}-values) for each isolate. Ej97 was mainly employed for this task, but K1 was additionally used on a few isolates to look at cross-resistance.

The desired bacteriocin start concentration was prepared by diluting 1 mg/ml bacteriocin stock solution with BHI medium, and kept on ice until further use. In this case, a start concentration of 200 µg/ml was utilized, implying a final diluted concentration of 50 µg/ml in the first well of each column, as shown in figure 3.2.

Materials:
5 ml indicator wild type ON-culture
5 ml resistant indicator mutants ON-culture
BHI medium
Ej97 or K1
Microtiter plate

Procedure:

1. Using a multi-pipette, 100 µl BHI medium was transferred to all 96 wells of the microtiter plate.
2. Next, 100 µl of 200 µg/ml bacteriocin was transferred to all the wells of column 1, before being two-fold diluted along each row by mixing and transfer of 100 µl along the rows (well 1-11). Row 12 was left as a negative control without any bacteriocin added
3. 100 µl 50x diluted (100 µl pure cultures in 5 ml BHI medium) of the various resistant isolates and wild types were added to all the wells of the row corresponding to the isolate.
4. Finally, the plates were mixed gently on a microtiter plate shaker and incubated at 30 °C for an appropriate amount of time, usually approximately 5-6 hours for 50x indicator dilution. The inhibition was subsequently assessed visually and/or by use of a microtiter plate scanner.

The bacteriocin concentrations were known (figure 3.2) in each column and enabled approximate calculations of MIC_{50}-values for each isolate from the results.

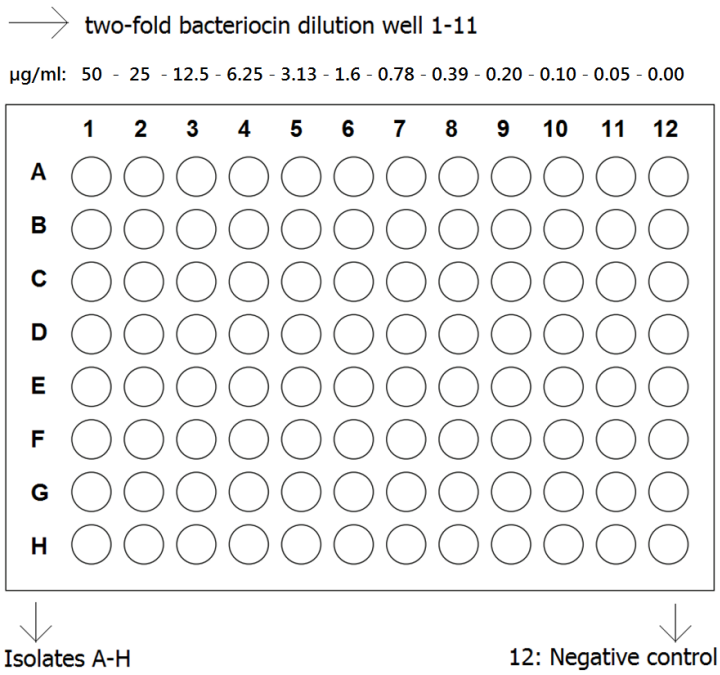


Figure 3.2: Diagram of the 96-well microtiter plates used in growth experiments for investigating phenotypes in relation to bacteriocin resistance. The bacteriocin was diluted two-fold along the wells of each row, but leaving row 12 as a negative control without any addition of bacteriocin. MIC_{50}-values could be estimated from the results.

3.6 DNA isolation

DNA from bacteriocin resistant mutants as well as the sensitive wild types was extracted from the same ON-cultures which had been applied to the microtiter plate assay. This would enable comparison of the genotypes to the observed phenotypes. The r*seP* was initially the point of interest due to being homologous with the gene encoding the LsbB receptor in *L. lactis*. As previously elaborated, the strong sequence similarity between LsbB and the three leaderless enterocins established RseP as the putative receptor for these bacteriocins as well.

To execute sequence analysis of *rseP* and ensure that the bacterial isolates did not derive from contaminants, DNA was extracted with a miniprep kit before *rseP* was amplified using species-specific *rseP* primers.

It was subsequently decided to initiate whole genome sequencing as several isolates which did not appear to have mutations within the putative receptor still exhibited a high level of resistance compared to the wild type. This required another protocol for DNA isolation involving a bacterial genomic DNA kit.

3.6.1 DNA isolation for *rseP* analysis

By combining FastPrep shaker with E.Z.N.A Plasmid MiniPrep kit, cell lysis and DNA extraction was accomplished through alkaline SDS lysis assisted by mechanical bead-beating, and by spin-columns specifically but reversibly binding DNA under optimized conditions. This allowed for contaminants and protein to be washed away before eluting the extracted DNA.

Materials:
Acid washed glass beads
Centrifuge
E.N.Z.A™ Plasmid MiniPrep Kit (section 2.7)
FastPrep homogenizer
FastPrep tubes
ON-culture

RNase A was added to solution I and absolute ethanol to the Wash Buffer as specified in the kit.

Procedure:

1. ON-culture was transferred to 1.5 ml eppendorf tubes and spun down at maximal speed (13 000g, 4 °C) for 4 minutes in order to yield an appropriately sized cell pellet. The supernatant was discarded.
2. The cell pellet was dissolved in 400 µl solution I before being transferred to a Fast-prep tube supplied with ~ 0.4 g of acid washed glass beads.
3. The Fast-prep tubes were secured properly in a Fast-prep homogenizer and run for 20 seconds at 6 m/s to.
4. The tubes were subsequently centrifuged at maximal speed for 5 minutes to enable collection of the supernatant without glass beads. The supernatant was transferred back into the proper eppendorf tubes.
5. 250 µl solution II was added and the eppendorf tube turned upside-down 4-6 times.
6. 350 µl solution III was added and the eppendorf tube onceover turned upside-down 4-6 for merging the components.
7. The tube was centrifuged at max speed for 10 minutes to enable removal of precipitates from the solution.
8. Subsequently the supernatant was transferred to Qiagen DNA binding spin columns and centrifuged for 1 minute at maximal speed before the flow-through was disposed of.
9. 500 µl of HB-buffer was added to the column and centrifuging for 1 minute before discarding the flow-through.
10. Step 9 was repeated using 750 µl wash buffer, before the sample was centrifuged for an additional 2 minutes.
11. With DNA still bound to it, the spin column was removed from the collection tube and placed in a new eppendorf tube.
12. DNA was eluted with 30-100 µl elution buffer depending on the initial cell pellet size. The elution buffer was carefully applied to the middle of the column and the tube was left open on the bench for 1 minute before being centrifuged at max speed for 1 minute.

The DNA quality and concentration were measured using NanoDrop before the samples were stored at -20 °C until further use. A minimal nucleic acid concentration of 50 ng/µl was desired, but the yield was usually well beyond this.

3.6.2 gDNA isolation for whole genome sequencing (WGS)

Preparation of genomic DNA for whole genome sequencing required a more thorough procedure with focus on purity and avoiding fragmentation of the gDNA. Here, complete lysis was achieved by a chaotropic salt-containing solution combined with a lysozyme for disrupting the thick peptidoglycan cell walls of the Gram-positive bacteria.

Materials:

ON-culture
Eppendorf centrifuge
Ethanol 96%
GenElute™ Bacterial Genomic DNA Kit (section 2.7)
Lysozyme L4919
Proteinase K
Water bath

Procedure:

Heating blocks were pre-heated to 37 °C and 55 °C respectively. Lysozyme solution was prepared by dissolving 40 mg lysozyme L4919 from chicken egg white in 1 ml sterilized water.

1. ON-culture was transferred to a 1.5 ml eppendorf tube and centrifuged at maximal speed for 4 minutes in order to yield an appropriately sized cell pellet. The supernatant was discarded.
2. The collected cell pellet was dissolved in 200 µl prepared lysozyme solution and incubated for 30 minutes at 37 °C to weaken the Gram-positive cell walls.
3. The incubation was followed by addition of 20 µl RNase A to remove RNA residues, and the tube was incubated at room temperature for 2 minutes.
4. The tube was subsequently supplemented with 20 µl proteinase K solution and 200 µl Lysis Solution C, before thorough mixing on vortex for ~15 seconds to achieve full homogenization of the components.
5. The sample was then incubated at 55 °C for 10 minutes to effectively lyse the cells.
6. Meanwhile, the DNA-binding column was prepared by applying 500 µl Column Preparation Solution prior to 1 min of centrifugation. The elute was discarded.

7. The tube was collected from the heating block and 200 µl of 96 % ethanol was added to the lysate and votexed for 15 seconds before the solution was carefully transferred to the DNA-binding column using a broad pipette tip to avoid gDNA fragmentation.
8. The column was centrifuged for 1 min and placed in a new collection tube.
9. 500 µl Wash Solution 1 was added before step 8 was repeated once more.
10. 500 µl Wash Solution Concentrate was then added and the columns centrifuged for a minimum of 3 minutes or until the column appeared completely dry.
11. The column was relocated to a clean eppendorf tube and the DNA eluted with 200 µl Elution Buffer. The Elution Buffer was carefully applied to the middle of the column and the tube was left open for 2 minutes before being centrifuged at max speed for 1 minute.

The gDNA was examined using Qubit fluorometer and agarose gel electrophoresis, and preserved at -20 °C until further use.

3.7 DNA measurements

3.7.1 NanoDrop

NanoDrop ND-1000 was used to obtain optical measurements of DNA-concentration and purity both after DNA isolation and PCR product clean-up. Elution buffer was employed for blanking the instrument, before 2 µl of DNA sample was applied to the pedestal.

The absorbance was measured at 260 nm as this is the absorbance maximum of nucleic acids. The concentration was returned in ng/ml while the purity was emitted by the instrument as the 260/280 absorbance ratio. This ratio should be within the range of 1.8 and 2.0. Lower values indicate contamination by proteins (absorbance maximum at 280 nm), and higher values indicate that RNA is present in the sample (e.g. when forgetting to add RNase to solution I).

3.7.2 Qubit

When working with gDNA for whole genome sequencing, a Qubit fluorometer was employed to evaluate the concentration of DNA. The Qubit in contrast to NanoDrop only measures DNA and is not influenced by protein contents or other contaminants.

Materials:

Qubit 2.0

0.5 ml Qubit tubes

Qubit dsDNA HS Kit Q32854 (section 2.7)

The working solution for the instrument was prepared by mixing 199 µl buffer and 1 µl dye per sample to be measured (including two standards). Next, the working solution was utilized in the preparation of standard 1 and 2 to make a standard curve for the Qubit instrument, which later was used to make sample calculations. 190 µl working solution and 10 µl standard were mixed in Qubit tubes using standard 1 and 2 respectively. DNA samples were prepared by thoroughly mixing 197 µl working solution and 3 µl eluted DNA in Qubit tubes before reading the concentration.

3.8 Primer construction and preparation

Primers for amplifying *rseP* had to be designed specifically for this thesis. Prior to primer design, the *rseP* sequence in each indicator species was retrieved using BLAST (Basic Local Alignment Search Tool) to search for homologues of the lactococcal *rseP (yvjB)*. The *yvjB* sequence (NC_002662.1, Lactococcus lactis subsp. lactis Il1403) was used as "input" for the algorithm, and each of the indicator species as "organism", respectively. The most significant hit was consistently a RIP metalloprotease RseP from the M50 superfamily. YvjB is the lactococcal RseP.

After locating the homologue in each species and retrieving the gene and its surrounding sequence, the flanking regions were examined to determine areas suitable as primers. The forward primers should be located at ~ 150 nucleotides (nt) upstream the gene, and the reverse primer ~ 100 nt downstream to ensure sequence data of good quality throughout the gene and the upstream regulatory region. The primer design resulted in theoretical PCR products of ~ 1600 nt. Middle primers approximately 600 nt into the theoretical PCR product were also designed to ensure complete data from the gene sequencing. The reverse primers were ordered as the reverse complement.

The success of the primers in a PCR reaction depends on several factors, many regarding the properties of the primers themselves. Satisfactory primer sequences should contain nucleotides with an even distribution of the different bases adenine (A), thymine (T), cytosine (C) and guanine (G). The 3' end base should be a C or T (the small pyrimidines), and the 3' should also preferably not contain stretches of weak bases (A and T). Additionally, the melting point should be between 64 °C and 70 °C which can be calculated easily by adding 2 °C for each A or T (double binding) and 4 °C for each C or G (triple binding). As a final control, the theoretical PCR products should be run through BLAST to ensure that the proposed primer regions are conserved across different strains of the species. This will substantiate that these areas likely are conserved in the strain the primers were designed for as well, ensuring optimal primer affinity.

All primers were ordered from Life Technologies and received in solid form. The mole amount (nmoles) for each primer was specified on the attached scheme and multiplied with 10 for calculating the dH_2O volume that was needed to obtain a 100 µM stock solution. Before being applied in PCRs, the primers were diluted once more by mixing 20 µl 100 µM primer solution and 180 µl dH_2O to get a 10 µM primer working solution.

3.9 Polymerase Chain Reaction (PCR)

Prior to sequencing, the *rseP* had to be amplified into larger quantities by PCR. Frequently utilized in many scientific fields, this is a well-known chain reaction driven by a heat-stable DNA polymerase acting on a DNA template and the components necessary for DNA replication in combination with alternating temperatures. Specific primers enable quick and simple production of many identical copies of desired DNA areas.

Materials:
Extracted template DNA (section 3.6.1)
10x PCR buffer
50 mM MgCl
10 mM dNTP
10 nM Forward primer (table 2.3)
10 nM Reverse primer (table 2.3)
Taq polymerase enzyme
dH_2O

Procedure:

While working on ice

1. The necessary amount of master mix was prepared was prepared in accordance to table 3.1 and made with a total volume sufficient for the number of PCR reactions needed (48 µl per sample). Taq Polymerase was the last component to be added.
2. The master mix was thoroughly merged and distributed into 0.2 ml PCR tubes.
3. 2 µl of template DNA was added to the correspondingly marked PCR tube.
4. The tubes were briefly spun down to remove bubbles and droplets from the walls/lid.
5. Finally, the tubes were placed in the PCR machine and run with the appropriate program according to the primers used for the reaction (table 3.2 and 3.3).

After the amplification process was complete, the samples could be stored at 4 °C for short periods until further use, or in a freezer.

Table 3.1: Volumes of the components used in the PCR reactions. Master mix preparation was done according to this table, but adding the templates after the master mix had been split into separate PCR tubes.

Reagent	Volume per tube
10x PCR buffer	5 µl
50 mM MgCl	1.5 µl
10 nM dNTP	1 µl
10 nM Forward primer	1 µl
10 nM Reverse primer	1 µl
Taq polymerase	0.5 µl
dH_2O	39 µl
template DNA	1 µl
Total volume	**50 µl**

Table 3.2: PCR program settings used in correlation with the *rseP* primers. After the initial denaturation of dsDNA to ssDNA, 30 cycles of amplification by denaturation, primer annealing and extension was executed before a final extension step and storage at 4 °C

Temperature	Time	Cycles	Action
95 °C	2 min	1	initial denaturation
95 °C	10 seconds		denaturation
63 °C	20 seconds	30	primer annealing
72 °C	1 min		primer extension
72 °C	5 min	1	final extension
4 °C	infinite		storage

Table 3.3: PCR program settings used in correlation with *rseP* primers for the *E. faecium* 2787 transposase-mutants. After the initial denaturation of dsDNA to ssDNA, 30 cycles of amplification by denaturation, primer annealing and extension was executed before a final extension step and storage at 4 °C

Temperature	Time	Cycles	Action
98 °C	1 min	1	initial denaturation
98 °C	10 seconds		denaturation
60 °C	25 seconds	30	primer annealing
72 °C	50 seconds		primer extension
72 °C	5 min	1	final extension
4 °C	infinite		storage

Table 3.4: PCR program settings specified in the protocol for MiSeq sample preparation in the WGS section 3.12.2, used in correlation with the index primers from the MiSeq Indexing Kit.

Temperature	Time	Cycles	Action
95 °C	30 seconds	1	initial denaturation
95 °C	10 seconds		denaturation
55 °C	30 seconds	12	primer annealing
72 °C	30 seconds		primer extension
72 °C	5 min	1	final extension
°C	infinite		storage

3.10 Agarose gel electrophoresis

PCR was followed by gel electrophoresis to confirm amplicon of the right size. Utilizing an agarose gel matrix and applying an electric force, the negatively charged DNA will move down the gel towards the positive pole. This process separates fragments according to size as smaller molecules will travel faster through the agarose matrix. By labeling the nucleic acids with the fluorophore GelRed, a substance structurally similar to ethidium bromide but less harmful, exposure to UV light will reveal how the bands have migrated in the gel. Having implemented a ladder with known band sizes in one well facilitates quick determination of sample band sizes.

1 % agarose was assembled for gel electrophoresis of *rseP*, and 0.7 % agarose for gDNA (larger fragments). This was accomplished by autoclaving agarose (2.5 g or 1.75 g respectively) in 250 ml 1xTAE. 1xTAE was prepared by diluting 50xTAE stock solution with distilled water. The agarose solutions were kept on 55 °C water bath between uses for gel-casting.

Materials:

1xTAE

Agarose

Parafilm

PCR product (section 3.9)

1kb ladder

GelRed

Loading buffer 6x

Gel electrophoresis equipment

GelViewer

Procedure:

1. The liquid agarose gel was poured into a gel molding form with a comb for making the wells, and allowed to solidify for 20-30 minutes.
2. The gel was removed from the cast and placed in the gel electrophoresis vessel where additional 1xTAE was poured over the gel until the wells were completely covered.
3. Then, the ladder and PCR products were mixed with loading buffer (LB) and GelRed on Parafilm according to table 3.5, and applied in the wells using a pipette.
4. The lid was subsequently attached to the vessel and plugged into the power source with approximately 90 V applied. The separation process was upheld until the bands had migrated a sufficient distance down the gel. This usually required about 30 minutes.
5. Finally, the gel was carefully removed from the vessel and studied using a GelViewer with UV-light to visualize the DNA bands.

Table 3.5: Components used for agarose gel electrophoresis were assembled according to this table. The components were merged as droplets on Parafilm to make the procedure less messy. The ladder was placed in the first well of the gel while the various amplicons were placed in the remaining wells next to it.

Sample:	Components in each well:
Ladder	5 µl 1 kb ladder (w/ LB) 1 µl GelRed
PCR product	5 µl amplicon 2 µl loading buffer (LB 6x) 1 µl GelRed

3.11 PCR product cleanup/purification

After verifying that the PCR reaction had resulted in a properly sized amplicon, the product was purified using a specialized kit to remove residues from the PCR reaction.

Materials:
PCR product (section 3.9)
Nucleospin® PCR Clean-up Gel Extraction kit
Eppendorf Centrifuge

Procedure:

1. 100 µl NT1 (a buffer facilitating binding of DNA to the silica membrane) was applied and mixed thoroughly into the PCR product, and the solution transferred to a spin column.
2. The solution was centrifuged at maximal speed for 1 minute.
3. Then, 700 µl of NT3 was applied on the column and centrifuged for 1 minute to wash away contaminants, before the collected fluid was discarded.
4. The column centrifuged for 2 additional minutes to dry the column completely before placing it in a fresh eppendorf tube.
5. 30 µl of elution buffer NE was added to the middle of the column and the column lid left open on the bench for 1 minute.
6. Finally, the column was centrifuged for 1 minute and the eluted DNA subsequently collected in the eppendorf tube was stored in a freezer for later use.

The quality of the purified DNA sample was measured using NanoDrop. The appropriate concentration for sequencing at GATC was 20-100 ng/ml.

3.12 DNA sequencing

The various collected DNA from mutant and wild type isolates were sequenced in order to acquire information for comparing the genotypes to the observed phenotypes. DNA variation could potentially explain the significant differences in bacteriocin sensitivity and why the bacterium no longer was recognized or harmed by the bacteriocin. Initially, the putative receptor

gene *rseP* from approximately 80 mutant isolates was sequenced to affirm that this LsbB-homologous gene had mutated and was responsible for causing resistance to Ej97 as well. Secondly, whole genome sequencing was carried out for a collection of 40 mutant isolates to shed light on why many resistant isolates did not have the expected *rseP* mutations.

3.12.1 *rseP* sequencing

GATC Biotech (http://www.gatc-biotech.com) was engaged for prompt and inexpensive single-gene sequencing of amplified and purified *rseP*. By utilizing the SUPREMErun sequencing service on their Sanger ABI 3730xl fleet, this technology provides longer single reads of up to 1,100 bp. Before sending the samples to GATC, mixtures consisting of 2.5 µl 10 µM primer and 7.5 µl purified amplicon (20-100 ng/ml) were merged in fresh eppendorf tubes. This step was repeated in three separate tubes for each DNA sample, applying the forward, middle or reverse primer respectively. GATC barcodes were attached to the tubes as means of sample identification when the sequence data was made available the GATC database.

3.12.2 Whole genome sequencing (WGS)

As explained in section 3.6.2, the GenElute protocol had been applied to collect gDNA of good quality and minimal fragmentation for this following procedure. The gDNA yield was confirmed on agarose gel and diluted to 0.2 ng/µl using Qubit measurements prior to MiSeq sample preparation. This would ensure that the samples met the optimal requirements for the MiSeq (5 µl sample with 1 ng DNA).

<u>The MiSeq instrument:</u>

The wild type and 10 mutants from each selected indicator strain (2 control isolates containing already identified *rseP* mutation, and 8 isolates where no mutation had been identified so far) were whole genome sequenced using the Nextera XT Sample Preparation Kit and the Illumina MiSeq sequencing instrument. Thus, a total of 44 samples would be sequenced in one run, which was within the MiSeq limits for maintaining sufficient coverage for a good assembly.

The MiSeq can yield ~25 million forward and reverse reads of approximately 300 bases each in one run (per flow cell). A coverage of ~50 per isolate is normally required for good assembly. This implies that the amount of genomes sequenced in one run should be equal to a maximum of:

Maximum genomes = total MiSeq output in bp / bp necessary per genome for good assembly

Maximum genomes = (25 000 000 reads×300 bp/read)/(genome size bp×50 coverage)

Maximum genomes = (25 000 000 reads×300 bp/read)/(3 000 000 bp×50 coverage)

Maximum genomes = (7 500 000 000 bp)/(150 000 000 bp)

Maximum genomes = 50

Using a genome size of 3 Mb (overestimate) in this calculation, 50 genomes can theoretically be sequenced, but due to the error rate it should be less than this. Here it was decided that 44 samples would most likely be within the MiSeq limitations as the genome sizes were also expected to be less than 3 Mb. Commonly, 75 % of the reads will have an adequate error rate of less than 0.001 per base (>Q30). However, this is dependent on an optimal clustering of 1200K/mm^2 to yield the necessary reads per flow cell, which again is affected by the DNA concentration and sample preparation.

Sample preparation for MiSeq sequencing run:

The Nextera XT DNA Sample Prep Kit was used to prepare the samples for whole genome sequencing. The preparation procedure would result in a paired end library for each sample before library pooling and sequencing was initiated. The preparation for MiSeq sequencing is relatively lengthily and is described in five steps. Consistent replacement of pipette tips is crucial to avoid cross-contamination throughout the MiSeq sample preparation.

Materials:

Agencourt AMPure XP 60 ml kit (beads) (section 2.7)
MiSeq Run Reagents Kit (V3-600 cycles) (section 2.7)
Nextera XT DNA Sample Prep Kit (96-samples) (section 2.7)
Nextera XT Indexing Kit (96-indeces) (section 2.7)
Nextera XT Plate Fixture kit (section 2.7)
96-well PCR plates
96-well microtiter plates
96-well thermal cycler with heated lid (PCR machine)

dH_2O

Ethanol absolute

Ethanol fresh 80 %

Heating block for 1.5 ml centrifuge tubes

Magnetic stand

Microplate centrifuge

Microplate shaker

Microseal for 96-well PCR plates

Tween 20

Step 1: Tagmentation

After measuring DNA concentration by Qubit and diluting the samples to 0.2 ng/µl, a process termed tagmentation was initiated. This step exploits an engineered transposome which couples the process of fragmenting the DNA with tagging the ends with unique adapter sequences. The MiSeq is capable of distinguishing these adapter sequences. Thus they provide a means of identification to differentiate reads from various isolates in the pooled sample. These unique adapter sequences are also the basis for primer annealing and initiating amplification of input DNA in the subsequent PCR (step 2).

Procedure:

1. ATM (Amplicon Tagment Mix), TD (Tagment DNA Buffer) and diluted input DNA from the freezer was thawed on ice while NT (Neutralize Tagment Buffer) was brought to room temperature from 4 °C storage. The reagents were gently mixed by inverting the tubes several times before being briefly spun down in a microcentrifuge.
2. A 96-well PCR plate was labeled according to the desired placement of the 44 DNA isolates. A 96-well microtiter plate was for support of the PCR plate.
3. Following this, 10 µl of TD, 5 µl ATM and 5 µl of input DNA was added to the corresponding wells.
4. The contents of the wells were gently mixed used a multi-channel pipette before the plate was covered with microseal and centrifuged in a microplate centrifuge at 280 g for 1 minute at 20 °C to collect all reagents at the bottom of the tubes.

5. Next, the plate was placed in a thermo cycler (PCR machine) for 5 minutes at 55 °C with the lid heated, followed by being held at 10 °C until the sample contents reached the same temperature.
6. 5 µl of NT was added to each well and gently mixed by pipetting to neutralize the tagmented samples before once again being covered with microseal and centrifuged for 1 minute at 280 g at 20 °C.
7. Finally, the plate was incubated at room temperature for 5 minutes.

Step 2: PCR Amplification

For amplifying the tagged and fragmented DNA templates into larger quantities, a limited-cycle PCR program was used in correlation with indexing primers 1 (N7) and 2 (S5). For 44 samples, indexing primers S501-07 and N701-07 were used (7x7=49) as described in figure 3.3.

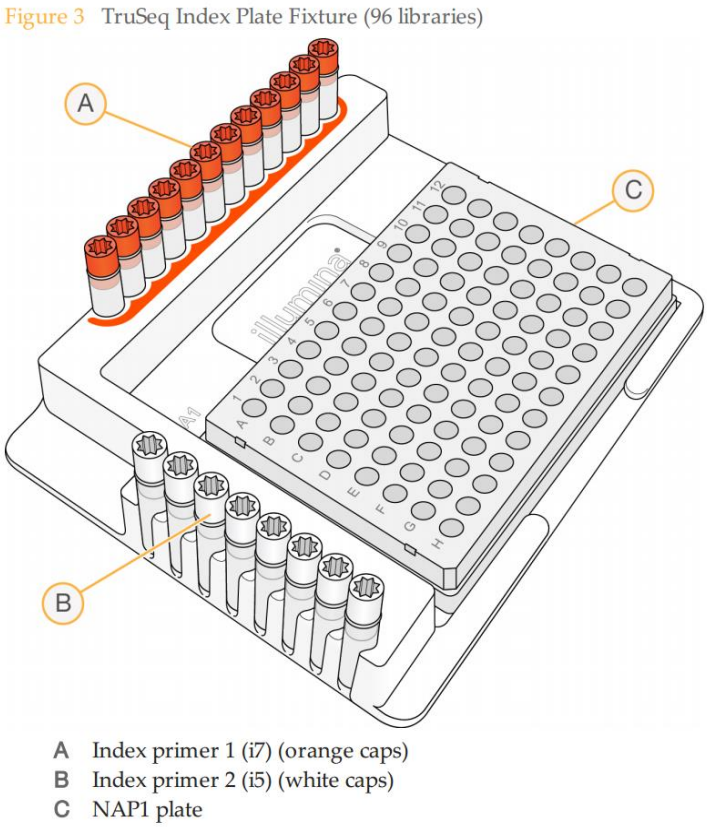


Figure 3.3: TruSeq Plate Fixture with a capacity of 96 samples, assembled with the primers of the indexing kit and the sample plate. For the 44 samples used in this experiment, indexing primers 1 (N701-N707) were arranged horizontally (orange caps), and indexing primers 2 (S501-S507) arranged vertically (white caps).

Procedure:

1. The indexing primers were thawed at room temperature and mixed briefly by inverting before being spun down in a microcentrifuge.
2. The primers were then arranged in the TruSeq Index Plate Fixture to simplify the mixing of reagents. Index 1 primers (N701-07) with orange caps were arranged horizontally and index 2 primers (S501-07) with white caps were arranged vertically (figure 3.3).
3. A new 96-well PCR plate (placed in a microtiter plate) was placed in the fixture and 15 µl from each of the wells of the old plate transferred to the new one. An additional 5 µl of the proper index primer 1 and 5 µl of the proper index primer 2 were then added to the wells. After the primers had been applied they were given fresh caps to minimize risk of contamination affecting later sample preparations.
4. The sample plate was covered with microseal and centrifuged at 280 g at 20 °C for 1 minute before PCR was initiated using the program settings specified in table 3.4.

Safe stopping point for cold storage at 2-8 °C for up to two days. Use microseal.

Step 3: PCR Clean-up

The DNA paired end library now having been attained by PCR was purified in using magnetic AMPure XP beads for removal of both very small DNA fragments and left-overs from the PCR.

Procedure:

1. The AMPure XP beads were brought to room temperature and fresh 80 % ethanol prepared.
2. The 96-well PCR plate containing amplicons from the previous PCR step was brought to room temperature and centrifuged for 1 minute to collect condensation from the microseal.
3. A fresh 96-well PCR plate was labeled as before and all the PCR products transferred from each well of the previous plate.
4. The AMPure XP beads were vortexed vigorously for 30 seconds to attain an even dispersion, before 30 µl was transferred to each well containing PCR product. The components were gently mixed by pipetting up and down 10 times.

5. The plate was next incubated at room temperature for 5 minutes before being placed at a magnetic stand for 2 minutes until the supernatant had completely cleared, and could be carefully discarded without collecting any beads (which were now associated with DNA).
6. While keeping the plate on the magnetic stand, the beads were washed using 200 µl 80 % ethanol without re-suspending the beads. After 30 seconds the supernatant was again discarded.
7. The wash-step was repeated once more before discarding the supernatant and letting the beads air-dry for 5 minutes on the magnetic stand.
8. After the beads had dried, 52.5 µl RSB (re-suspension buffer) was added to each well and gently mixed with the beads by pipetting 10 times to bring the DNA from the beads to the solution.
9. The plate was incubated at room temperature for 2 minutes before once more being placed on the magnetic stand.
10. A new 96-well PCR plate was labeled as before and 50 µl transferred from each of the wells of the previous plate (amplified and purified tagged-end libraries) without including any beads.

    Safe stopping point for cold storage -15 to -20 °C for up to one week. Use microseal.

Step 4: Library Normalization

To ensure a more equal representation of each sample before pooling the libraries, the quantity of each library was normalized using normalization beads.

Procedure:

Reagents used in this step contain formamide which is a probable reproductive toxin, and should be disposed of accordingly. Avoid inhalation, ingestion, and skin/eye contact.

1. LNA1 (Library Normalization Additives 1) was removed form the freezer, as well as LNB1 (Library Normalization Beads 1), LNS1 (Library Normalization Storage Buffer 1) and LNW1 (Library Normalization Wash 1) from the fridge, and brought to room temperature.

2. A new 96-well PCR plate was labeled as before and 20 µl was transferred from the previously used sample plate with wells containing purified PCR products.
3. 2.2 ml LNA1 was added to a clean 15 ml tube.
4. LNB1, containing the beads, was vortexed for 1 minute to ensure complete suspension of the bead pellet, before 400 µl immediately was transferred to the tube containing LNA1.
5. 45 µl of the LNA1/LNB1 solution was quickly added to each of the 44 wells containing 20 µl DNA sample.
6. The plate was subsequently sealed with microseal and placed securely on a microplate shaker at 1800 rpm for 30 minutes for the DNA fragments to attach and normalize on the bead surfaces.
7. After the incubation period, the plate was again put on a magnetic stand for 2 minutes to draw the beads with DNA attached from the solution. After the supernatant had cleared, it was carefully discarded in an appropriate hazardous container using a pipette whilst on the magnetic stand.
8. The plate was removed from the magnetic stand and the beads washed by adding 45 µl LNW1 to each well before sealing the plate with microseal and putting the plate back on the microplate shaker at 1800 rpm for 5 additional minutes.
9. Next, the plate was once more placed on the magnetic stand for 2 minutes and the supernatant with contaminants discarded.
10. Step 8 and 9 was repeated one more time to ensure sufficient washing of the beads.
11. After the beads were thoroughly washed, the DNA fragments attached to them could be eluted. 30 µl of 1 M freshly prepared NaOH was added to each well before the plate was sealed with microseal.
12. The contents were merged on the microplate shaker for 5 minutes at 1800 rpm.
13. The plate was subsequently placed on the magnetic stand for 2 minutes to draw the beads from the supernatant now containing eluted DNA libraries in normalized amounts.
14. A new 96-well PCR plate was labeled and 30 µl LNS1 added to each well, followed by 30 µl with eluted DNA from the wells of the previous plate.
15. Finally, the plate was sealed with microseal and centrifuged at 1000 g for 1 minute.

Safe stopping point for cold storage using microseal.

Optional step - Preparation of PhiX control:

In this case it was decided to add more diversity to the sample libraries to enhance the confidence in the sequencing run. This was attained by adding PhiX control. The PhiX is a control library consisting of fragments the MiSeq is capable of recognizing and differentiating from the samples. E.g. in 16S rRNA sequencing, the fragments are very similar to each other and it is common to add ~30% of PhiX. However, in whole genome sequencing using DNA libraries from various isolates and species, the fragments have much less similar sequences, making it unnecessary to add PhiX in such large quantities. PhiX increases diversity among the sequence fragments and can improve confidence in the MiSeq output. Also, it helps estimate error rates and determines if errors are likely due to sample preparation mistakes or the MiSeq.

Procedure:

1. 2 µl 10 nM PhiX library and 3 µl 10 mM Tris pH 8.5 (RSB) was combined in a fresh eppendorf tube to attain a 4 nM PhiX library.
2. Following this, 5 µl of the 4 nM PhiX library and 5 µl freshly prepared (within the last 12 hours) 1 M NaOH was combined in a new tube and vortexed briefly, before the sample was spun down at 280 xg for 1 minute, attaining a 2 nM PhiX library. The contents were incubated for 5 minutes at room temperature to denature the DNA of the PhiX control.
3. Finally, 990 µl chilled HT1 was added to the 10 µl PhiX dilution.

The 20 pM PhiX control can be stored for up to 3 weeks between -15 and -20 °C before use.

Step 5: Library pooling and MiSeq sample loading

The MiSeq uses a pooled sample consisting of equal volumes of normalized sample libraries, which is diluted in hybridization buffer and heat denatured before sequencing can be engaged.

Procedure:

1. A heating block compatible with 1.5 ml eppendorf tubes was pre-heated to 96 °C, and an ice bath prepared by combining 3 parts ice and 1 part water.
2. The MiSeq reagent cartridge was removed from freezer storage and thawed at room temperature.

3. The 96-well sample plate containing purified and amplified DNA libraries was also thawed at room temperature and centrifuged at 1000 g at 20 °C for 1 minute to collect condensation.
4. The contents of each well were mixed gently by pipetting before 5 µl from each well were combined in a clean eppendorf tube (pooled amplicon library, PAL).
5. As a final step, 570 µl HT1, 5 µl PhiX and 25 µl PAL were mixed in a new eppendorf tube labeled DAL (diluted amplicon library), and the tube was mixed thoroughly by vortexing at top speed.
6. The DAL tube was subsequently incubated for 2 minutes at 96 °C on the heating block.
7. The DAL tube was immediately after incubation inverted 1-2 times and placed on the ice bath for 5 minutes to attain denaturation.
8. Finally, the denatured and pooled sample could be loaded into the MiSeq reagent cartridge. The cartridge was placed in designated enclosure in the MiSeq which would now run for a couple of days to complete sequencing of the pooled library. Sample information was registered in the Illumina Experiment manager.

MiSeq washing routines

After each run on the MiSeq, post-run washes are carried out using Tween 20. First, 5 ml 100 % Tween 20 is added to 45 ml of dH_2O to attain a 10 % Tween 20 solution. Then, 25 ml of the 10 % Tween 20 is added to 475 ml dH_2O constituting the MiSeq was solution. 350 ml of this 0.5 % Tween 20 wash solution is poured into the wash bottle and "Post-run wash" selected from the MiSeq interface.

Other washing procedures are also done regularly to uphold optimal performance of the MiSeq, and include maintenance wash once a month and stand-by wash when the MiSeq will not be used for the next 7 days.

3.13 Analysis of sequence data

Various tools were implemented in the analysis process of the *rseP* and moreover in the whole genome sequences. Numerous software and tools are available to choose from, and they can give significantly different results based on the algorithm implemented in them. This demands a critical approach to analysis of sequence data.

rseP analysis - pipeline

After downloading the data sets from the first sequencing procedure (which revolved around merely the rseP), various bioinformatics tools were used to identify mutations present in the putative receptor. In the GATC database, each sequenced isolate was accompanied by three relatively small datasets of AB1-files (each resulting from sequencing affiliated with forward, middle or reverse primer, as previously mentioned) which ensured a simplistic analysis process.

1. BioEdit (installed from http://bioedit.software.informer.com) was employed to visualize the sequence signals/base calls stored in the AB1-files, and to evaluate the quality of the data (figure 3.4). Sequence strings of low ambiguity were extracted as plain text and copied to a regular text editor. Tools implemented in BioEdit could also have been used to complete the analysis but in this case it was decided to preform the assembly manually.
2. The three text data sets could easily be combined into the full gene sequence by identifying overlapping areas. E.g. the extracted sequence from dataset E1_F (forward primer), E1_M (middle primer) and E1_R (reverse primer) were merged into E1 (the complete *rseP* sequence of *E. faecalis* 3358 mutant isolate E1).
3. Following the assembly of the *rseP* for each sample, ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2) provided a tool for multiple sequence alignments (MSAs) of the data, and thus elucidating areas differentiating mutants form wild type. Default presets were set for excecuting the algorithm.
4. NCBI ORF finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) was finally used to establish the correct reading frame (figure 3.5) of the *rseP*-genes and easily identify and visualize the mutational consequences on the translated protein. Mutations resulting in premature stop codons or frame shifts were especially interesting to locate as they were likely to result in a fully inactivated bacteriocin receptor.

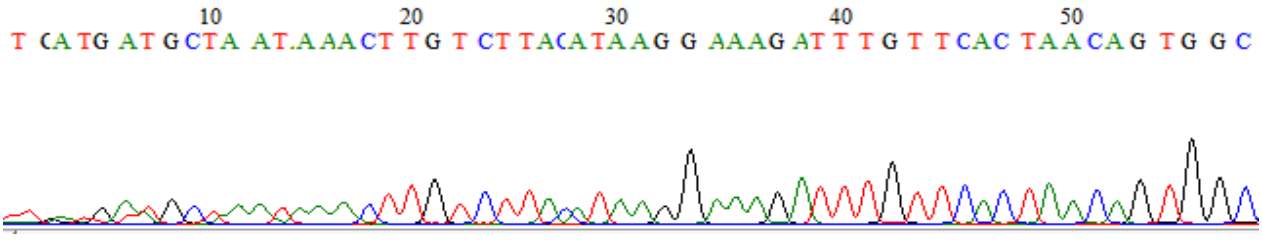


Figure 3.4: Sequence stretch from GATC with base call signals/chromatogram visualized in BioEdit. The base calls at the beginning (first ~30 bp) are of poor quality (ambiguous bases) and were removed prior to assembly.

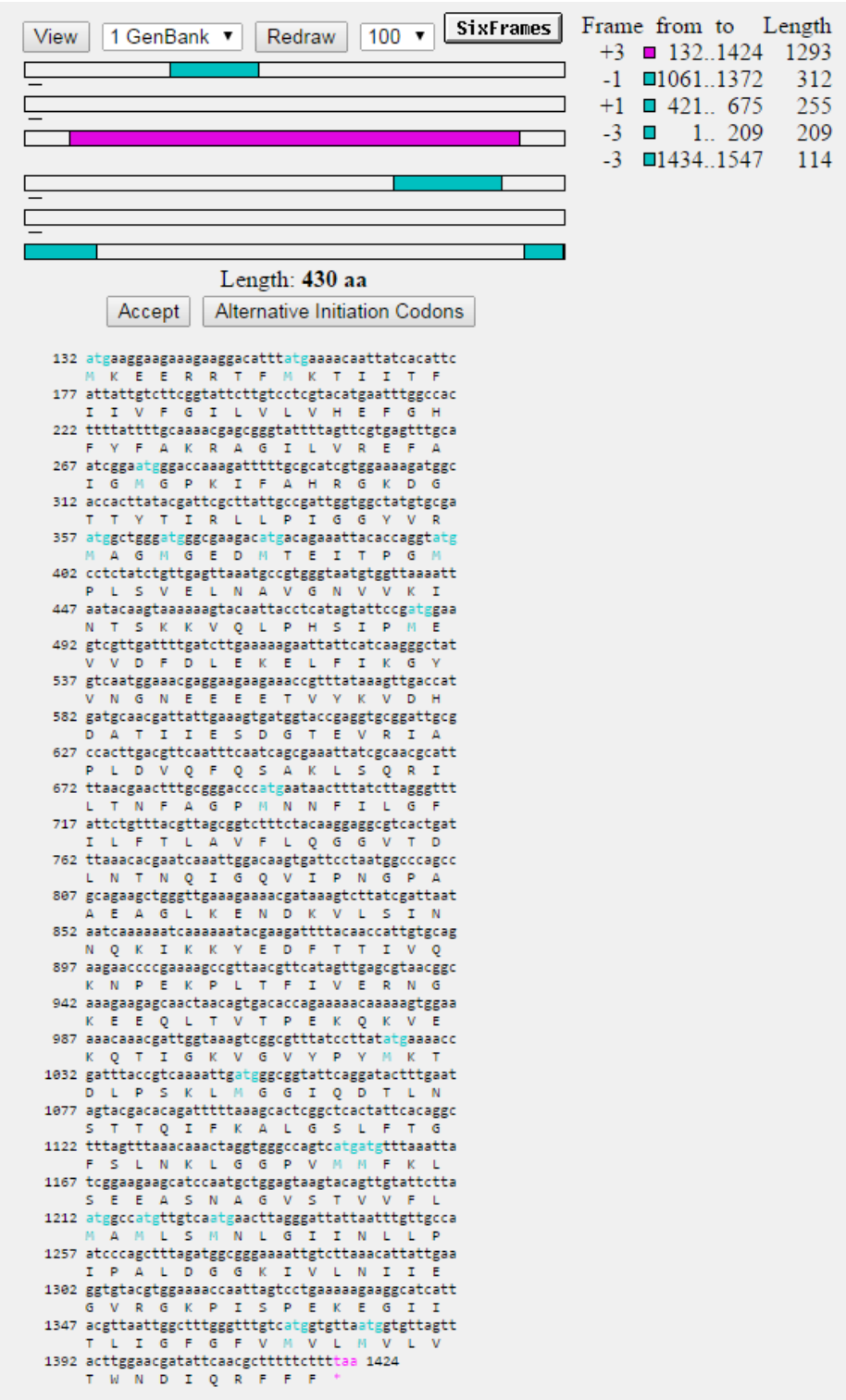


Figure 3.5: Output from NCBI ORF finder when using the assembled *rseP* sequence from *E. faecalis* 3358 wild type as input. This picture establishes frame +3 as the correct reading frame (the only alternative resulting in a protein of the proper length). Using *rseP* sequence from the mutants as input could easily reveal how DNA mutations resulted in altered gene products.

WGS analysis – Pipeline

This section describes a simple method on how to analyze sequence data from whole bacterial genomes. The data presented here were mainly processed in CLC genomics workbench 7.5/5.5 (http://www.clcbio.com), but many other tools, could rather have been employed instead.

1. Fastq-files from the MiSeq output files were made available in Illumina BaseSpace (https://basespace.illumina.com/home/index) and imported to CLC Workbench. Every sample had two fastq-files affiliated with it, termed R1 and R2 respectively. Each of these files consists of several hundred thousand reads of approximately 300 nt, as well as quality data which is basically what separates fastq files from the common fasta files.
2. The reads were initially run through a QC quality analysis to evaluate lengths distribution and PHRED quality scores, among other parameters (figure 3.6). This step can provide valuable information but is not essential to the data analysis process. *CLC Toolbox > NGS core tools > Create sequencing QC report.*

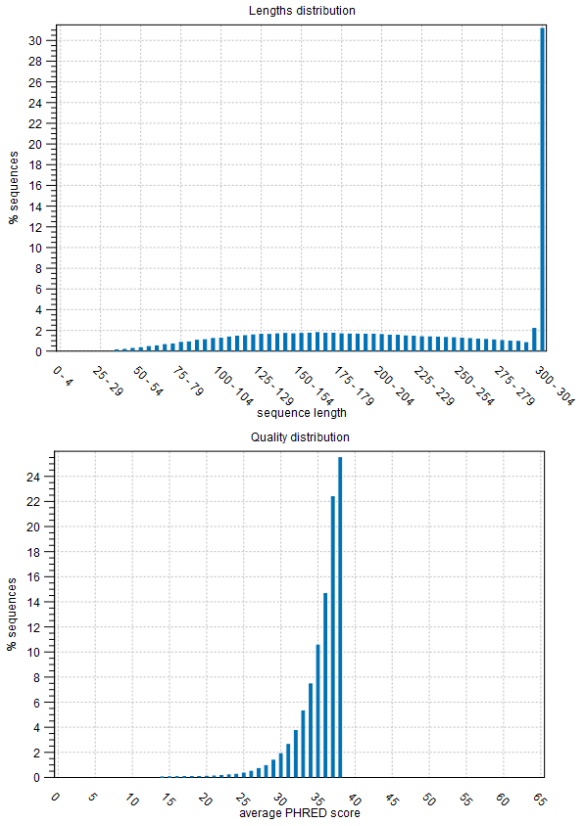


Figure 3.6: The lengths distribution of reads (on top) and the PHRED quality scores (at bottom) after QC analysis of *E. faecalis* 3358 wild type data.

3. The first task to accomplish was assembly of the four wild type genomes. The remaining 40 mutants would be handled later. All reads were trimmed using default settings, except a quality score limit of 0.001 and minimum length 50 nt in order to discard reads of poor quality. *CLC Toolbox > NGS core tools > Trim sequences*.
4. Next, de novo assembly was initiated with the algorithm implemented in CLC. This would assemble reads into longer contigs by identification of overlapping sequence data. Contigs below 1000 bp were set to be discarded. This step is often sufficient for variant detection, and step 5 and 6 can thus be excessive. *CLC Toolbox > De novo sequencing > De novo assembly.*
5. To establish the right order of the assembled contigs of the wild types, the genome alignment visualizer software Mauve 3.2.1 (downloaded from http://gel.ahabs.wisc.edu/mauve/download.php) was employed. The de novo assembly output from CLC was submitted as input in the Mauve contig mover together with a similar reference genome. Default settings were used. After preforming the alignment of the contigs to the reference genome in several iterations to attain an optimal result, the output was visualized (figure 3.7) and re-imported to CLC as a fasta alignment file.

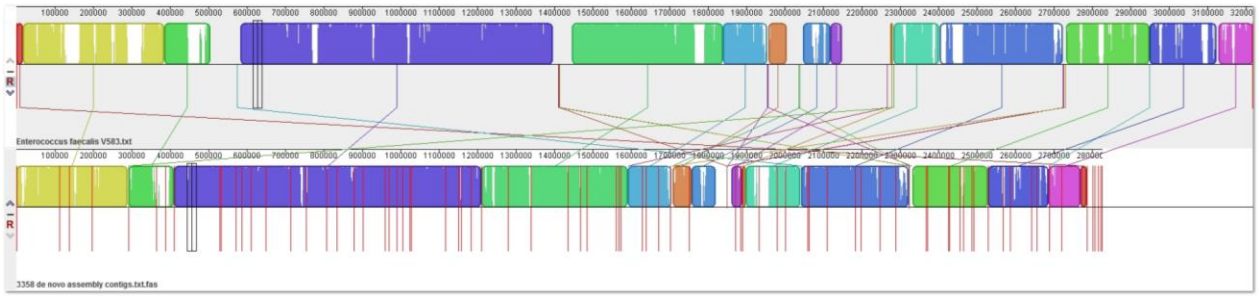


Figure 3.7: Output after 4 iterations of aligning the contigs of *E. faecalis* 3358 wild type to the reference sequence NC_004668.1 in mauve contig sorter. The reference genome is shown on top and the draft genome on the bottom.

6. After Mauve contig mover had established the right arrangement of the contigs, concatenation was executed to finish the draft genome. Although this would not result in a fully correct sequence, it was sufficient for use in this work. *CLC toolbox > Alignments and trees > Join alignments (concatenation).* However, this did not work as smoothly as expected due to CLC importing sequences in one single file. Instead the concatenation

was executed manually using Geneious 7.1.5 to import the Mauve alignment file with the sequences kept separate. There, the contig sequences could be ordered in a list according to the positions specified by Mauve, before being exported as a single new text file where gaps were removed, and the sequences manually joined/concatenated. This was a very ineffective approach and concatenation could have been omitted from the pipeline in this case. For the remaining steps, the de novo assembled contigs from step 4 were employed as input instead.

7. To annotate the wild type contigs, RAST (http://rast.nmpdr.org) or Prokka (http://www.vicbioinformatics.com/software.prokka.shtml) were used to locate gene elements in the genome and import them to CLC as GenBank files as seen in figure 3.8.

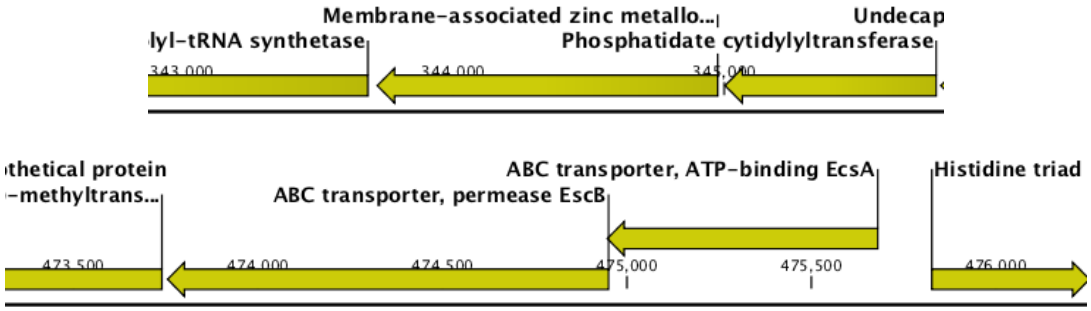


Figure 3.8: The locations of a few genes in *P. acidilactici* 2002 as predicted by RAST annotation and displayed in CLC workbench. The RAST GenBank output file was downloaded and imported to CLC workbench.

8. The next task was to compare the sequence data from the mutants to the wild types. First, the mutant reads were trimmed as previously done with the wild types.
9. The reads from the mutants were subsequently mapped separately to the wild type contigs (now termed reference genome) to uncover differences. *CLC Toolbox > NGS core tools > map reads to reference*.
10. Coverage analysis was carried out to look for abnormally low or high read coverage along the reference genome, which could indicate i.e. large deletions. *CLC Toolbox > Resequencing analysis > Coverage analysis.*
11. Concluding the genome analysis pipeline, probabilistic variant detection was executed to look for smaller variants to elucidate SNPs (single nucleotide polymorphism) and MNVs (multiple nucleotide variation) common across in many mutants and possibly strains. These could hold a potential of being involved in development of Ej97-resistance. A

workflow was created for this (figure 3.9). *CLC Toolbox* > *Resequencing analysis* > *Probabilistic variant detection.* Tips: If a candidate gene prevails quickly, annotate the upstream and downstream areas of that gene before proceeding with the variant detection. This can help locate possible mutation outside the coding region/in the regulatory area of the gene in other mutants and strains that would otherwise be more difficult to notice as significant.

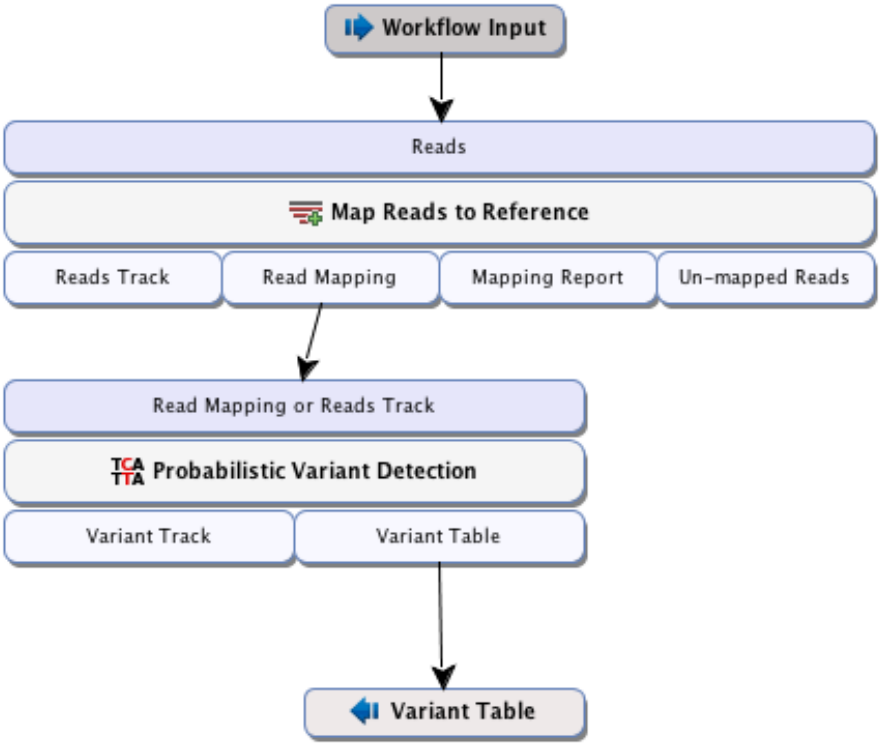


Figure 3.9: Workflow created for locating variants in CLC workbench using a probabilistic variant caller. The mutant reads were used as inputs and first mapped to the reference wild type contigs. This output of the mappings was subsequently used as input in the probabilistic variant detection. The output was given as variant tables showing SNPs and MNVs. The algorithm detects he probability for each combination of alleles is calculated. These probabilities are then used to determine which one of the allele combinations is the most likely combination for each position, and do not detect low frequency alleles. (http://www.clcsupport.com/clcgenomicsworkbench/current/index.php?manual=Probabilistic_variant_detection.html).

After the Probabilistic variant detection, the candidate genes were extracted from each of the isolates and aligned to the wild type in order to display the mutational impacts on DNA sequence and aa-sequence (after translation by CLC).

3.14 Unspecific resistance assay

With the analysis of all the sequence data completed, the mutants lacking alterations within *rseP* were screened for unspecific resistance mechanisms. This was carried out by employing the unrelated bacteriocins BHT-B and Garvicin-ML. If the isolates displayed increased resistance to these as well as to Ej97, it could be a good indication of the presence of a general resistance mechanism in the mutants.

BHT-B is another leaderless bacteriocin, but is not related to Ej97 and produced by *Streptococcus rattus*. Garvicin-ML is a circular bacteriocin which is folded in a manner which shields it from being downgraded by peptidases, perhaps increasing its spectrum of activity. The inaccessibility of recognition sites as a consequence of a tightly folded three-dimensional structure is likely responsible for this resistance to proteolytic enzymes (Borrero et al. 2011). For instance, if a mutation causes increase secretion of peptidases, the linear leaderless bacteriocins will likely all be degraded and unable to exert their activity, whilst the circular Garvicin ML will remain active. In the cross-resistance assay, the theory was that presence of an unspecific resistance mechanism will cause the mutants to be resistant to other unrelated bacteriocins in addition Ej97, but perhaps not to Garvicin ML due to the previously elaborated arguments.

Microtiter plate assays were used for the purpose of investigating this possibility. A 200 µg/ml start concentration of Ej97 and the unrelated bacteriocin BHT-B were used to screen a selection of mutants for similar resistance patterns. The wild types, as well as mutants known to have putative receptor mutations, were also included in the assay. The plates were incubated for approximately five hours at 30 °C and read visually.

An active fraction of Garvicin ML was also used of a selection of isolates, but the precise concentration was not known in this case. However, it was merely the patterns of resistance levels compared to Ej97-resistance that was of particular interest.

3.15 Mutation frequency

In the case of *E. faecalis* 3358, certain interesting *rseP*-mutations were detected in the form of CAAAAAAT-repeat variance, and subsequently 150 new mutants were generated for additional experiments using the same technique described in section 3.4.

The objective was to investigate the frequency of each repeat-variant by using mutation-specific primers. New primers were designed specifically for this purpose (table 2.3, figure 3.10) and used in PCR to differentiate between the numbers of repeats. After PCR, the amplicons were applied in agarose electrophoresis, and thus presence of DNA bands when observed through the GelViewer would indicate that the isolate had the number of repeats (genotype) corresponding to the primer, without the need for sequencing.

Unfortunately, the experiment was not successful, as only DNA smears or presence of products where they theoretically not should be, were observed. The experiment was thus not finished at this point in time as PCR with the designed primers did not yield the expected amplicons when using control DNA with know repeat number, and it was decided to rather pursue other tasks.

```
Del_68: CAAAAAT deletion (1 repeat)                    GTCTTATCGATTAATAATCAAAAAATAC
3358wt: Wild type (2 repeats)                        CGATTAATAATCAAAAAATCAAAAAATAC
Ins_68: CAAAAAT insertion (3 repeats)              TAATCAAAAAATCAAAAAATCAAAAAATAC

>3358 WT rseP gene
GGCGAAGTGGTCAAGTCCAATGGTTATTTATGAGTGAATATGCTAAAATAAAACTTGTCTTACATAAGGAAAGATTTGTTCACTAA
CAGTGGCGACTGTTCTCAGTAAACAGTTGCACTGGTTTATCGGGAAGGGAACAGAGCAATTTTTACGAGACTTTCCCATGTAATAA
TAAAAAGCACTGAACGTCAGTGAATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGTATT
CTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTAGTTCGTGAGTTTGCAATCGGAATGGGACC
AAAGATTTTTGCGCATCGTGGAAAAGATGGCACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGA
TGGGCGAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGTAATGTGGTTAAAATTAATACA
AGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAAGTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGT
CAATGGAAACGAGGAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACCGAGGTGCGGATTG
CGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATTTTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTA
GGGTTTATTCTGTTTACGTTAGCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATTCCTAA
TGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAATAATCAAAAAATCAAAAAATACGAAGATTTTA
CAACCATTGTGCAGAAGAACCCCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGACACCA
GAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATATGAAAACCGATTTACCGTCAAAATTGATGGG
CGGTATTCAGGATACTTTGAATAGTACGACACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAG
GTGGGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTGTATTCTTAATGGCCATGTTGTCA
ATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAGCTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGG
AAAACCAATTAGTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGGTGTTAGTTACTTGGA
ACGATATTCAACGCTTTTTCTTTTAAGGAGAAAGCTTAAGTAAAAAAGTGTGTAAAGAAATCAAAAATAGTTGATTTTAAAATAGA
GGAGAGTTTTAATGAAACAGTCAAAAATGTTAATTCCAACGTTAAGAGAAGTACCAAATGATGCAGAAGTCTTAAGTCATC
```

Figure 3.10: The mutation specific forward primers designed for each of the repeat variants in the frequency experiment on top, and below the *E. faecalis* 3358 *rseP* gene sequence with repeat area (yellow), reverse primer (green) and desired amplicon between these (~350 bp).

4. RESULTS

4.1 Inhibition spectrum assay

A total of 52 indicator bacteria confirmed distinct activity differences between the leaderless bacteriocins LsbB, Ej97 and K1, as listed in table 4.1. Ej97 and K1 displayed a broad spectrum with Ej97 having slightly wider activity than K1, while LsbB demonstrated activity towards certain *L. lactis* strains.

Nisin was employed as a control due to its well-defined activity spectrum, which includes a broad range of Gram-positive bacteria due to targeting the common peptidoglycan cell-wall precursor lipid II. Nisin is not active by itself to Gram-negative species. However, while the assay confirmed the broad activity of nisin, activity was also documented towards an *E. coli* strain (*E. coli* LMG 3235), which suggest contamination during the experiment or error in the glycerol stock. The same *E. coli* strain was also sensitive to LsbB which substantiates that some kind of mistake must have been made.

Table 4.1: The results from the spot-on-lawn inhibition spectrum assay. Various strains from the LMG glycerol stock had been employed in assessing the activity range of Ej97, K1 and LsbB. Nisin with its well-defined activity was implemented in the experiment as a control. "+" indicates that an inhibition zone was observed, while "++"occasionally has been used to highlight significant size-differences of the zones.

Indicator strain	Nisin	Ej97	K1	LsbB
E. coli LMG 2746	+	+	+	
E. coli LMG 3235	+	+	+	+
Enterobacter LMG 3287	+	+		
B. cereus LMG 2805				
L. innocua LMG 2710	+	+	+	
L. innocua LMG 2785	+	+	+	
L. ivanovii LMG 2813	+	+	+	
L. monocytogenes LMG 2604	+	+	+	
L. monocytogenes LMG 2650	+	+	+	
L. monocytogenes LMG 2651	+	+	+	
L. monocytogenes LMG 2652	+	+	+	
L. monocytogenes LMG 2653	+	+	+	

S. aureus LMG 3022				
S. aureus LMG 3023	+			
S. aureus LMG 3242	+	+	+	
S. salivarius LMG 1301	+			
L. sakei LMG 2334	+			
L. sakei LMG 2356	+	+	+	
L. sakei LMG 2361	+	+		
L. sakei LMG 2380				
L. sakei LMG 2799	+			
L. curvatus LMG 2353				
L. curvatus LMG 2715	+	++	+	
L. curvatus LMG 2355				
L. curvatus LMG 2371	+			
L. plantarum LMG 2003	+	+	+	
L. plantarum LMG 2352	+			
L. plantarum LMG 2357	+			
L. plantarum LMG 2358	+			
L. plantarum LMG 2362	+			
L. plantarum LMG 2379	+			
L. plantarum LMG 3125	+			
P. acidilactici LMG 2002	+	++	+	
P. pentosaceus LMG 2001	+	++	+	
P. pentosaceus LMG 2366	+	+	+	
P. pentosaceus LMG 2722	+	++	+	
L. lactis LMG 2081	+			
L. lactis LMG 2705	+	++	+	+
L. lactis LMG 2130	+	+	+	+
L. lactis LMG 3419	+	+		
L. strain F4-13 LMG 2070	+	++	+	
L. garvieae DCC43 LMG 3390	+	+	+	
E. avium LMG 3465	+	+		
E. faecium LMG 2787	+	+	+	
E. faecium LMG 2763	+			
E. faecium LMG 2783		+		
E. faecium LMG 2876		+		
E. faecium LMG 2787	+	+	+	
E. faecalis LMG 2333	+	+		
E. faecalis LMGT 3358	+	++	+	
E. faecalis SMF37 LMGT 3370	+	++	+	
E. faecalis DEC23 LMGT 3386	+	++	+	
L. gelidium LMG 2386	+			

Two of the indicators used in the assay simultaneously produced naturally resistant mutants during the experiment. These strains were *E. faecalis* 3358 and *P. acidilactici* 2002 (figure 4.1), and they were to be the basis of further work progress.

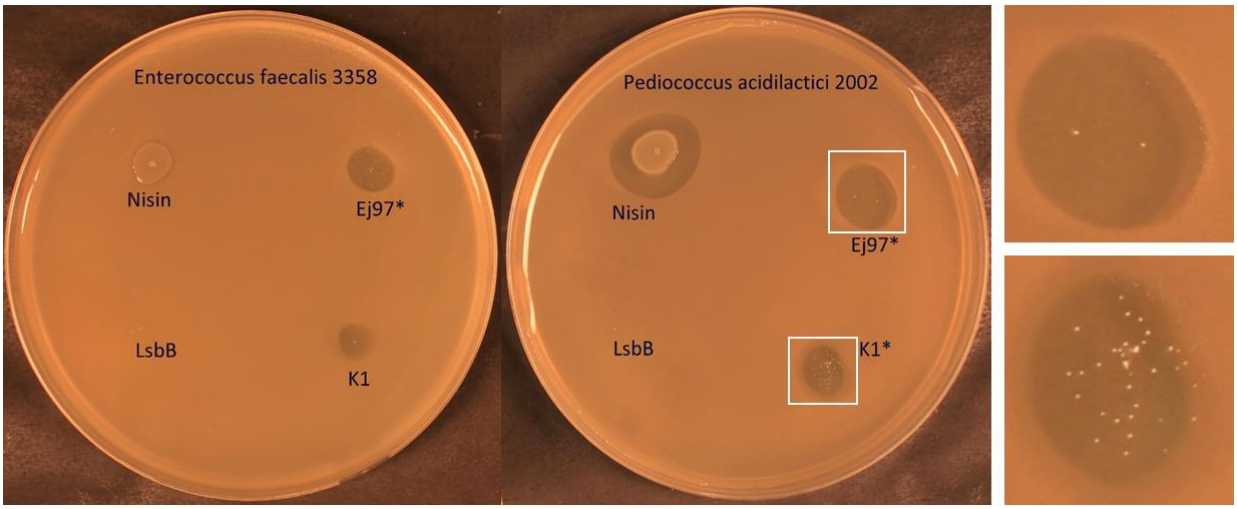


Figure 4.1: Pictures of the plates containing indicators producing the initial mutants for this study, and which arose naturally when exposed to 0.5 mg/ml bacteriocin. Resistant colonies can be observed in the zones marked with *. Close-ups of the *P. acidilactici* Ej97- and K1-resistant mutants are inserted to the far right.

4.2 Accumulation of bacteriocin resistant mutants

The two indicators *E. faecalis* 3358 and *P. acidilactici* 2002 which initially resulted in emergence of resistant colonies were subsequently exposed to various bacteriocin concentrations to produce more resistant colonies. Mutants from three other strains (*E. faecium* 2787, *L. garvieae* DCC43 and *L. garvieae* 1546) were later supplemented by other lab members. This yielded a large collection of mutants which were isolated and stored as glycerol stock (table 4.2).

Table 4.2: The collection of resistant mutants produced for this study. Only mutants that were later sequenced are included in this table. CFU = colony forming unit (within bacteriocin inhibition zone).

Indicator bacteria:	Name of isolate:	CFU picked from inhibition zone (mg/ml):
E. faecalis LMG 3358	3358 WT	
	E1	Ej97 0,5
	E2	Ej97 0,5
	E3	Ej97 0,25
	E4	Ej97 0,25

	E7	Ej97 0,25
	E8	Ej97 0,1
	E10	Ej97 0,5
	E11	Ej97 0,5
	E12	Ej97 0,25
	E14	Ej97 0,25
	E15	Ej97 0,1
	E17	Ej97 0,1
	E20	Ej97 0,5
	E22	Ej97 0,5
	E23	Ej97 0,25
	E24	Ej97 0,25
	E25	Ej97 0,25
	E26	Ej97 0,1
	E27	Ej97 0,1
	E28	Ej97 0,1
P. acidilactici LMG 2002	2002 WT	
	M1	Ej97 0,5
	M2	Ej97 0,5
	M3	Ej97 0,5
	M4	Ej97 0,25
	M5	Ej97 0,25
	M6	Ej97 0,25
	M7	Ej97 0,1
	M8	Ej97 0,1
	M9	Ej97 0,1
	M10	K1 0,5
	M11	K1 0,5
	M12	K1 0,5
	M13	Ej97 0,5
	M14	Ej97 0,5
	M15	Ej97 0,5
	M16	Ej97 0,5
	M17	Ej97 0,5
	M18	Ej97 0,25
	M19	Ej97 0,25
	M20	Ej97 0,25
	M21	Ej97 0,25
	M22	Ej97 0,25
L. garvieae LMG DCC43	DCC43 WT	
	G1	Ej97 0,5
	G2	Ej97 0,5
	G3	Ej97 0,5
	G4	Ej97 0,5
	G5	Ej97 0,5
	G6	K1 0,5
	G7	K1 0,5
	G8	K1 0,5
	G8	K1 0,5
	G10	K1 0,5

L. garvieae LMG 1546	1546 WT	
	G11	Ej97 0,5
	G12	Ej97 0,5
	G13	Ej97 0,5
	G14	Ej97 0,5
	G15	Ej97 0,5
	G16	K1 0,5
	G17	K1 0,5
	G18	K1 0,5
	G19	K1 0,5
	G20	K1 0,5
E. faecium LMG 2787	2787 WT	
	B28	EntQ 1,0
	B30	EntQ 1,0
	B32	EntQ 1,0
	B36	EntQ 1,0
	B38	EntQ 1,0
	B42	EntQ 1,0
	B44	EntQ 1,0
	B48	EntQ 1,0
	B70	K1 0,5
	B72	K1 0,5
	B74	K1 0,5
	B76	K1 0,5
	B80	K1 0,5
	B84	K1 0,5
	B86	K1 0,5
	B94	K1 0,5

4.3 Microtiter plate assays

The bacteriocin resistance levels were estimated in the mutants and compared to the wild types by using microtiter plates with two-fold dilutions of Ej97 (table 4.3). The results demonstrated that mutants from *E. faecalis* 3358 and *P. acidilactici* 2002 typically fell into one of two subgroups based on this phenotype. The first group appeared to be completely resistant even at concentrations above 50 µM, and the second group exhibited a lower level of resistance but was still many-fold more tolerate to bacteriocin than the wild type.

The same pattern could not be detected for the remaining three strains, *E. faecium* 2787, *L. garvieae* DCC43 and *L. garvieae* 1546, at the concentrations applied (50 µM being the highest concentration), but neither did the results disprove the possibility of these groupings being present in these instances as well. Corresponding assays using higher bacteriocin concentrations would provide a definitive answer to this.

Table 4.3: Microtiter assay results describing the resistance levels of the various isolates as MIC_{50}-values.

Isolate:	Ej97 MIC_{50} value (µg/ml)
3358 WT	0.4
E1	12.5
E2	12.5
E3	>50
E4	50
E7	25
E8	50
E10	>50
E11	12.5
E12	12.5
E14	25
E15	50
E17	>50
E20	>50
E22	>50
E23	25
E24	25
E25	>50
E26	25
E27	25
E28	>50
2002 WT	0.2
M1	>50
M2	6.25
M3	3.125
M4	3.125
M5	>50
M6	>50
M7	12.5
M8	12.5
M9	6.25
M10	25
M11	12.5
M12	6.25
M13	3.125
M14	3.125
M15	12.5
M16	3.125
M17	>50
M18	>50
M19	6.25
M20	>50
M21	>50
M22	6.25
DCC43 WT	0.8
G1	>50
G2	>50
G3	>50
G4	>50

G5	>50
G6	>50
G7	>50
G8	>50
G9	>50
G10	>50
1546 WT	6.25
G11	>50
G12	>50
G13	>50
G14	>50
G15	>50
G16	6.25
G17	>50
G18	>50
G19	>50
G20	>50
2787 WT	0.6
B28	>50
B30	>50
B32	>50
B36	>50
B38	>50
B42	>50
B44	>50
B48	>50
B70	>50
B72	>50
B74	>50
B76	>50
B80	>50
B84	>50
B86	>50
B94	>50

4.4 *rseP* sequence analysis

The findings within *rseP* corresponded perfectly with the groupings of resistance levels observed in *E. faecalis* 3358 and *P. acidilactici* 2002. Isolates which exhibited exceptionally high resistance in the microtiter assay consistently also contained serious mutations within this gene or in close proximity upstream it. As discussed in the previous section, this pattern could not be confirmed for the three other indicators. The *rseP* mutations and mutational impacts on RseP are presented as multiple sequence alignments (MSAs) in the appendix (A1-A8).While many isolates indeed contained the expected *rseP* mutations, the majority had intact *rseP* and thus this analysis could not provide explanations for the observed resistance.

E. faecalis 3358: Different genotypes were present in this strain as repeat variance of an 8 bp sequence stretch "CAAAAAAT", located near the middle of the gene. Whilst the wild type contains two such repeats, a deletion or insertion induces frameshift and premature termination in certain mutants. Of the 20 sequenced isolates, only 7 had a mutated *rseP* in the form of this addition or deletion of one repeat (table 4.4).

P. acidilactici 2002: Diverse and scattered *rseP* mutations, including an up-stream gene mutation, frameshifts, missense mutations, and nonsense mutations, constituted the genotypes of this strain. Of the 20 sequenced mutants, 10 isolates had mutation in or upstream *rseP* (table 4.4).

L. garvieae DCC43: Nonsense mutations or framehifts resulting in premature termination of the RseP were identified in the *rseP* of all the 10 isolates sequenced from this strain (table 4.4).

L. garvieae 1546: In the case of this *L. garvieae* strain, only 3 of the 10 sequenced isolates had mutations within (nonsense mutations) or upstream the putative receptor gene *rseP* (table 4.4).

E. faecium 2787: The DNA fragments attained from PCR indicated that some of the mutants had transposons inserted into the gene, making the amplicon appear as approximately double size on agarose gel (figure 4.5). Hence it was necessary to design additional primers to obtain coverage the whole gene sequence. After acquiring the complete sequence data it was ascertained that transposons were present in the mutated *rseP*.

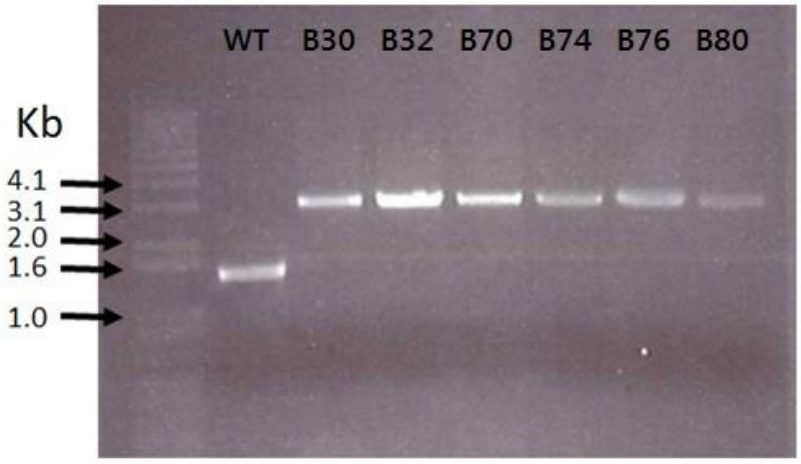


Figure 4.5: The gel picture illustrates the increased size of *rseP* in certain *E. faecium* 2787 mutants. On the left is the ladder with band sizes stated, followed by the *E. faecium* 2787 wild type *rseP* amplicon at ~1.6 kb, and finally six mutant amplicons (B30, B32, B70, B74, B76 and B80) which indicate a gene size of ~3000 bp.

All the sequenced mutants from this particular strain had either a transposon inserted in *rseP*, or no mutation at all within this gene. Additionally, all the transposons were found to be inserted at the very beginning of the gene sequence, or even before the start codon in the upstream gene region (figure 4.6). Isolate B70 and B32 had transposons inserted in the upstream region of the gene, while B76, B80, B84, B94, B30 and B74 had transposons shortly after the initiation codon. "X" marks the placement of the various transposons, all in near proximity of the start codon marked in blue.

```
2787WT      ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATG--------G AAAA
B28         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATG--------G AAAA
B36         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATG--------G AAAA
B38         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATG--------G AAAA
B42         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATG--------G AAAA
B44         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATG--------G AAAA
B48         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATG--------G AAAA
B72         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATG--------G AAAA
B86         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATG--------G AAAA
B32         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCTXAAAAAGCTATG AAAA
B80         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATGAAAACGAT AAAA
B74         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATG--------G AAAA
B76         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATG--------GXAAAA
B84         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATGAAAACGAT AAAA
B30         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATG--------G AAAA
B70         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAGGAGAAAAAXGAGAAAAAAGCT ATG--------G AAAA
B94         ACGTTGTTTAAAAAAACACACAAAAAACAAGTAAAG-------- GAGAAAAAAGCT ATG--------G AAAA
            ************************************        ************ * *        ****

2787WT      CGAT ACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B28         CGAT ACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B36         CGAT ACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B38         CGAT ACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B42         CGAT ACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B44         CGAT ACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B48         CGAT ACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B72         CGAT ACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B86         CGAT ACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B32         CGAT ACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B80         CGATXACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B74         CGAT ACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTTXTGATTGTTCATGAGTTTGGTCATTTCT
B76         CGAT ACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B84         CGATXACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B30         CGAT ACTAACATTTATCATTXGTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B70         CGAT ACTAACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
B94         CGATX----ACATTTATCATT GTTTTTGGTATATTAGTGATTGTT --------CATGAGTTTGGTCATTTCT
            ****     ************ ************************         ******************
```

Figure 4.6: Placements of the transposons in each of the *E. faecium* 2787-mutants with transposons inserted. X marks the placement of the various transposons relative to the *rseP* start codon ATG, meaning that the position of an "X" in fact symbols a ~1300 bp transposon sequence. Short deletions or insertions of repeats from the original gene often surround the location where transposons have been inserted, and explain the grey sequence stretches present in the alignment. This alignment only show the first part of the *rseP* sequence. * marks conserved positions.

BLAST was employed for identifying homologues of the transposon sequences, and showed that all these ~1300 bp insertions were close to identical to transposons already identified in other *E. faecium* strains. Additionally, significant similarities between the transposons were uncovered by MSA in ClustalW. The phylogenetic tree of the alignment showed that the transposons grouped into subgroups according to similarity (figure 4.7). In the *E. faecium* 2787 strain, 8 of the 16 sequenced isolates contained a mutated *rseP* in this manner (table 4.4).

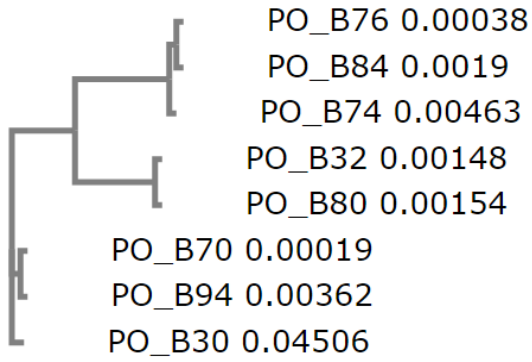


Figure 4.7: Phylogenetic tree displaying the similarities of the transposon sequences identified in *E. faecium* 2787 after MSA in ClustalW. The tree represents the proposed evolutionary history between the transposons, and indicates based on this prediction that certain transposons are very similar to each other.

The MSA of the transposon sequences are included in the appendix (A9).

Table 4.4: *rseP* sequencing results are listed below, with the locations of DNA and amino acid changes described. The default RseP lengths for these strains are between 375 to 431 aa. "*" indicates the position of termination. "-"denotes where *rseP* was found intact in the mutant.

Isolate:	Mutation type	*rseP* coding region change:	RseP amino acid change:
E1	-	-	-
E2	-	-	-
E3	Frameshift	732_739del CAAAAAAT	Ile244fs, Thr264*
E4	-	-	-
E7	-	-	-
E8	-	-	-
E10	Frameshift	740_747ins CAAAAAAT	Tyr247fs, Phe265*
E11	-	-	-
E12	-	-	-
E14	-	-	-
E15	-	-	-
E17	Frameshift	732_739del CAAAAAAT	Ile244fs, Thr264*
E20	Frameshift	740_747ins CAAAAAAT	Tyr247fs, Phe265*

E22	Frameshift	732_739del CAAAAAAT	Ile244fs, Thr264*
E23	-	-	-
E24	-	-	-
E25	Frameshift	740_747ins CAAAAAAT	Tyr247fs, Phe265*
E26	-	-	-
E27	-	-	-
E28	Frameshift	732_739del CAAAAAAT	Ile244fs, Thr264*
M1	Missense and Frameshift	833A>G, 835delG	Lys278fs, Met281*
M2	Missense	512G>A	Arg171Lys
M3	-	-	-
M4	-	-	-
M5	Nonsense	1021C>T,	Gln341*
M6	Frameshift	486delG	Gln162fs, Leu173*
M7	-	-	-
M8	-	-	-
M9	-	-	-
M10	Missense	512G>A, 1070C>A	Arg171Lys, Asn357Lys
M11	-	-	-
M12	-	-	-
M13	-	-	-
M14	-	-	-
M15	Missense	512G>A	Arg171Lys
M16	-	-	-
M17	Upstream gene	(9 bp up-stream gene)G>A	-
M18	Nonsense	184G>T,	Gly42*
M19	-	-	-
M20	Nonsense	506G>A,	Trp169*
M21	Nonsense	184G>T,	Gly42*
M22	-	-	-
G1	Frameshift	846delG	Pro281fs, Asp287*
G2	Nonsense	460C>T	Gln152*
G3	Frameshift	847insG	Pro281fs, Tyr286*
G4	Frameshift	846delG	Pro281fs, Asp287*
G5	Nonsense	718C>T	Gln238*
G6	Nonsense	502C>T	Gln166*
G7	Nonsense	355C>T	Gln119*
G8	Frameshift	847insG	Pro281fs, Tyr286*
G9	Frameshift	847insG	Pro281fs, Tyr286*
G10	Frameshift	846delG	Pro281fs, Asp287*
G11	-	-	-
G12	Upstream gene	(2 bp upstream gene)G>A	-
G13	-	-	-
G14	-	-	-
G15	-	-	-
G16	-	-	-
G17	Nonsense	584G>A	Trp195*

G18	-	-	-
G19	-	-	-
G20	Nonsense	328G>T	Glu110*
B28	-	-	-
B30	Transposon	~1300 bp inserted	Frameshift, premature stop
B32	Transposon	~1300 bp inserted	Frameshift, premature stop
B36	-	-	-
B38	-	-	-
B42	-	-	-
B44	-	-	-
B48	-	-	-
B70	Transposon	~1300 bp inserted	Frameshift, premature stop
B72	-	-	-
B74	Transposon	~1300 bp inserted	Frameshift, premature stop
B76	Transposon	~1300 bp inserted	Frameshift, premature stop
B80	Transposon	~1300 bp inserted	Frameshift, premature stop
B84	Transposon	-	-
B86	-	-	-
B94	Transposon	~1300 bp inserted	Frameshift, premature stop

4.5 Whole genome sequence analysis

The whole genome sequencing was carried out with a total of 44 isolates (table 4.5) to investigate the origin of the medium resistance level observed in the microtiter plate assays. CLC workbench and various other bioinformatics tools were used for analysis of the MiSeq data as elaborated in section 3.13.

Table 4.5: This isolates from the four strains that were subjected to WGS. Along with the wild type, 8 isolates without *rseP* mutations and 2 control isolates with identified *rseP* mutations were used from each strain. *L. garvieae* DCC43 was not included in WGS as all sequenced isolates already had shown *rseP* mutations.

Strain	*E. faecalis* 3358	*P. acidilactici* 2002	*E. faecium* 2787	*L. garvieae* 1546
Wild type:	wt3358	wt2002	wt2787	wt1546
Resistant isolates without *rseP*-mutation:	E2	M4	B28	G11
	E4	M7	B36	G13
	E8	M8	B38	G15
	E11	M11	B42	G14
	E12	M12	B44	G16
	E14	M16	B48	G18
	E24	M19	B72	G19
	E26	M22	B86	
Isolates with *rseP* -mutation:	E3	M1	B80	G12
	E25	M5	B84	G17
				G20

Coverage analysis was executed in CLC after mapping the mutant reads to the wild type contigs. However, the coverage analysis did not yield any distinctive or interesting results for either of the strains, and was subsequently followed up by a probabilistic variant detection to uncover smaller SNPs (single nucleotide polymorphism) and MNV (multiple nucleotide variation) in the mutants.

The non-processed probabilistic variant detection results are present in appendix A21-A23, but were trimmed and altered prior to presentation in table 4.6 to enable easier identification of variants present in several resistant isolates. Mutations appearing within many resistant bacteria and across species would likely be involved in the resistance observed, rather than being random mutations. The trimming of the results was done by assessing variant location, coverage, and frequency. Mutations close to the contig ends where the sequence data are more ambiguous, as well as hits with low read coverage and variants present in low frequencies were removed from the variant tables. In this case variants located outside annotated genes were also viewed as less relevant and removed if not present in high frequency.

The variant detection uncovered a pattern across three of the strains (*E. faecalis* 3358, *P. acidilactici* 2002 and *L. garvieae* 1546) which suggested involvement of an ABC transporter and its components EcsA and EcsB. Mutations were identified within these genes in all but one (isolate E4) of the mutants that did not have *rseP* mutation (table 4.6a-c), while they were absent from isolates which did have *rseP* mutations.

The *E. faecium* 2787 sequence data were not sufficiently good to attain a satisfactory contig assembly, which also caused the variant detection to result in around 300 hits per isolate (data not shown), most of them present in low frequencies and/or with low coverage. This suggested that the sequencing preparation was not as successful in this case as for the rest of the samples. It was subsequently decided not to invest more time in solving this matter. However, by focusing only on the EcsAB area of *E. faecium* 2787 (after the discovery having been made in the other strains), mutations were found in three of the mutants (data not shown); one mutation inside the *ecsA* (B72) and two mutations upstream the *ecsA* (B38 and B80). B80 additionally had transposon inserted in *rseP,* making it the only isolate in this experiment with mutations in both *rseP* and *ecsAB.*

Table 4.6: Results from probabilistic variant detection in a) *E. faecalis* 3358, b) *P. acidilactici* 2002, c) *L. garvieae* 1546 using CLC workbench 5. Hits were removed if present near edges of contigs, in low frequency, with low coverage, or outside genes. Complete and unedited results are shown in appendix A21-A23.

a) *E. faecalis* 3358

Isolate	Mapping	Contig length	Reference Position	Consensus Position	Variant type	Reference	Allele variants	Frequencies	Coverage	Overlapping annotations
E2	Contig_5	201796	141178	141141	SNV	C	A	100	33	CDS: ecsA
E3	Contig_13	92290	28928	28923	SNV	A	C	97.4	39	CDS: PROKKA_00689
	Contig_2	305519	5691	5691	InDel	CAAAAAAT	-	93.2	74	CDS: rseP
E4	-	-	-	-		-	-	-	-	-
E8	Contig_5	201796	141090	141053	InDel	A	-/A	55.6/44.4	18	CDS: ecsA
E11	Contig_5	201796	141178	141141	SNV	C	A	100	23	CDS: ecsA
E12	Contig_5	201796	141178	141141	SNV	C	A	100	17	CDS: ecsA
E14	Contig_15	78035	28172	28172	SNV	C	A/C	66.7/33.3	60	CDS: yjlD
	Contig_5	201796	141178	141141	SNV	C	A	100	73	CDS: ecsA
E24	Contig_1	360546	168503	168503	SNV	C	A	98.4	62	CDS: bglH
	Contig_13	92290	67706	67701	SNV	A	G	100	57	CDS: luxO
	Contig_13	92290	67973	67968	SNV	G	T	100	65	CDS: luxO
	Contig_2	305519	100859	100859	SNV	A	C	100	43	CDS: PROKKA_01228
	Contig_2	305519	108105	108105	SNV	G	A	100	69	CDS: mepM
	Contig_2	305519	271471	271469	SNV	T	C	100	31	CDS: hisS
	Contig_22	15523	9074	9074	SNV	A	C	98.5	67	CDS: PROKKA_01482
	Contig_3	255904	209677	209677	SNV	G	T	100	43	CDS: PROKKA_01754
	Contig_3	255904	221921	221921	SNV	C	T	100	68	CDS: PROKKA_01764
	Contig_5	201796	140041	140004	InDel	-	T	98.6	70	CDS: ecsB
E25	Contig_1	360546	168503	168503	SNV	C	A	98.7	77	CDS: bglH
	Contig_13	92290	67706	67701	SNV	A	G	100	56	CDS: luxO
	Contig_13	92290	67973	67968	SNV	G	T	100	82	CDS: luxO
	Contig_2	305519	5699	5699	InDel	-	CAAAAAAT	83.8	80	CDS: rseP
	Contig_2	305519	100859	100867	SNV	A	C	98.5	67	CDS: PROKKA_01228
	Contig_2	305519	108105	108113	SNV	G	A	100	86	CDS: mepM
	Contig_2	305519	271471	271477	SNV	T	C	100	50	CDS: hisS
	Contig_22	15523	9074	9074	SNV	A	C	98.9	90	CDS: PROKKA_01482
	Contig_3	255904	209677	209677	SNV	G	T	98.3	60	CDS: PROKKA_01754
	Contig_3	255904	221921	221921	SNV	C	T	100	70	CDS: PROKKA_01764
E26	Contig_5	201796	141178	141141	SNV	C	A	100	46	CDS: ecsA

b) *P. acidilactici* 2002

Isolate	Mapping	Contig length	Reference Position	Consensus Position	Variant type	Reference	Allele variants	Frequencies	Coverage	Overlapping annotations
M5	contig_6	598825	343959	343974	SNV	G	A	100	92	CDS: rseP
M1	contig_6	598825	344153	344168	InDel	T	-	100	81	CDS: rseP
M16	contig_6	598825	473899	473914	InDel	-	T	100	54	CDS: ecsB
M4	contig_6	598825	473899	473914	InDel	-	T	96.9	64	CDS: ecsB
M12	contig_6	598825	474610	474625	InDel	C	-	100	43	CDS: ecsB
	contig_1	392950	11716	11716	SNV	T	T/G	76.5/23.5	115	CDS: T1SS
M19	contig_6	598825	474847	474862	InDel	C	-	100	22	CDS: ecsB
M11	contig_6	598825	474847	474862	InDel	C	-	100	26	CDS: ecsB
	contig_1	392950	11716	11716	SNV	T	T/G	72.5/27.5	80	CDS: T1SS
	contig_1	392950	11718	11718	SNV	T	T/G	74.7/25.3	83	CDS: T1SS
	contig_1	392950	11724	11724	SNV	C	C/A	75.3/24.7	73	CDS: T1SS
	contig_1	392950	11728	11728	SNV	A	A/G	73.2/26.8	71	CDS: T1SS
	contig_1	392950	11732	11732	SNV	G	G/C	76.0/24.0	75	CDS: T1SS

M7	contig_6	598825	474836	472837	SNV	T	C	100	12	CDS: ecsB
M22	contig_6	598825	475043	475058	SNV	C	T	100	37	CDS: ecsA
	ontig_6	598825	475145	475160	InDel	T	-	96.9	32	CDS: ecsA
	contig_1	392950	11716	11716	SNV	T	T/G	70.7/29.3	92	CDS: T1SS
	contig_1	392950	11718	11718	SNV	T	T/G	73.2/26.8	97	CDS: T1SS
	contig_1	392950	11724	11724	SNV	C	C/A	72.8/27.2	92	CDS: T1SS
	contig_1	392950	11728	11728	SNV	A	A/G	72.1/27.9	86	CDS: T1SS
	contig_1	392950	11732	11732	SNV	G	G/C	74.2/25.8	89	CDS: T1SS
	contig_1	392950	11737	11737	SNV	G	G/C	73.6/26.4	87	CDS: T1SS
	contig_1	392950	11746	11746	InDel	C	C/-	73.5/26.5	83	CDS: T1SS
	contig_1	392950	11752	11752	MNV	TT	TT/CG	74.5/22.7	141	CDS: T1SS
	contig_1	392950	11755	11755	SNV	T	T/C	74.4/25.6	82	CDS: T1SS
M8	contig_5	48938	19081	19081	SNV	C	A	100	75	CDS: conjugation
	contig_6	598825	474315	474330	MNV	GCC	T	97.1	210	CDS: ecsB
	ontig_1	392950	11716	11716	SNV	T	T/G	73.0/23.0	152	CDS: T1SS
	contig_1	392950	11718	11718	SNV	T	T/G	75.8/24.2	153	CDS: T1SS

c) *L. garvieae* 1546

Isolate	Mapping	Contig length	Reference Position	Consensus Position	Variant type	Reference	Allele variants	Frequencies	Coverage	Overlapping annotations
G13	Contig_2	225134	105252	105252	SNV	C	T	100	31	CDS: ecsB
	Contig_8	94072	91239	91239	InDel	G	-	79.5	39	CDS: PROKKA_02168
G11	Contig_2	225134	104573	104573	SNV	G	T	100	46	CDS: ecsA
	Contig_8	94072	91239	91239	InDel	G	-	94.8	77	CDS: PROKKA_02168
	Contig_9	86285	41078	41078	InDel	A	-/A	83.3/16.7	12	
G12	Contig_17	27701	9109	9109	SNV	C	T	100	51	CDS: rseP
	Contig_19	12279	12244	12244	SNV	G	A	95.2	21	
	Contig_8	94072	91239	91239	InDel	G	-/G	75.0/25.0	52	CDS: PROKKA_02168
	Contig_9	86285	41078	41078	InDel	A	-/A	70.0/25.0	20	
G20	Contig_17	27701	8780	8780	SNV	C	A	100	57	CDS: rseP
	Contig_8	94072	91239	91239	InDel	G	-/G	72.1/26.5	68	CDS: PROKKA_02168
	Contig_9	86285	41078	41078	InDel	A	-	100	10	
G19	Contig_12	54019	53136	53136	SNV	G	C	99.2	126	CDS: PROKKA_00822
	Contig_2	225134	104874	104874	SNV	G	T	100	95	CDS: ecsA
	Contig_8	94072	91239	91239	InDel	G	-	83.6	61	CDS: PROKKA_02168
G18	Contig_2	225134	105493	105493	SNV	G	A	100	55	CDS: ecsB
	Contig_20	14878	257	185	SNV	A	A/C	64.8/35.2	71	CDS: PROKKA_01273
	Contig_8	94072	91239	91239	InDel	G	-	86.9	84	CDS: PROKKA_02168
	Contig_9	86285	41078	41078	InDel	A	-/A	50.0/50.0	16	
G14	Contig_1	685987	361233	361233	SNV	G	T	100	84	CDS: PROKKA_00385
	Contig_19	12279	7663	7663	SNV	C	C/A	68.6/31.4	236	CDS: PROKKA_01037
	Contig_2	225134	104851	104851	SNV	C	A	100	55	CDS: ecsA
	Contig_8	94072	91239	91239	InDel	G	-/G	74.4/25.6	39	CDS: PROKKA_02168
	Contig_9	86285	41077	41077	MNV	AA	A/-	52.5/32.5	40	
G15	Contig_2	225134	104558	104558	SNV	C	T	100	60	CDS: ecsA
	Contig_8	94072	91239	91239	InDel	G	-	82.3	62	CDS: PROKKA_02168
	Contig_9	86285	41078	41078	InDel	A	-	78.6	14	
G17	Contig_17	27701	8524	8524	SNV	C	T	100	112	CDS: rseP
	Contig_3	234027	222261	222261	SNV	C	T	98.2	55	CDS: PROKKA_01566
	Contig_8	94072	91239	91239	InDel	G	-	81.3	75	CDS: PROKKA_02168
	Contig_9	86285	41078	41078	InDel	A	-	93.8	16	
G16	Contig_2	225134	104749	104749	SNV	C	A	100	93	CDS: ecsA
	Contig_8	94072	91239	91239	InDel	G	-	82.6	86	CDS: PROKKA_02168

Green = Mutated *rseP*-gene

Blue = Mutated Ecs ABC transporter component *ecsA* (ATPase) or *ecsB* (permease).

After identifying these two ABC-transporter genes as likely being connected to resistance, the mutations were studied in more detailed and are described in table 4.7a-c. MSA of the EcsAB sequences showing location and impact of mutations are present in A15-A20.

Table 4.7: *rseP* and *ecsAB* mutational impacts in the a) *E. faecalis* 3358, b) *P. acidilactici* 2002, c) *L. garvieae* 1546 isolates implemented in WGS. Only for isolate E4 were no mutation in either gene identified. The *rseP*-mutations coincide with those of table 4.4 although the positions provided by CLC are somewhat different. "*" indicates the position of termination.

a) *E. faecalis* 3358

Isolate	Overlapping annotations	Coding region change	Amino acid change
E2	CDS: ecsA	c.382G>T	p.Glu128*
E3	CDS: rseP	c.708_715delCAAAAAAT	p.Ile236fs
E4			
E8	CDS: ecsA	c.[470delT];[=]	p.[Ile157fs];[=]
E11	CDS: ecsA	c.382G>T	p.Glu128*
E12	CDS: ecsA	c.382G>T	p.Glu128*
E14	CDS: ecsA	c.382G>T	p.Glu128*
E24	CDS: ecsB	c.780_781insA	p.Lys261fs
E25	CDS: rseP	c.715_716insCAAAAAAT	p.Tyr239fs
E26	CDS: ecsA	c.382G>T	p.Glu128*

b) *P. acidilactici* 2002

Isolate	Overlapping annotations	Coding region change	Amino acid change
M1	CDS: rseP	c.827delA	p.Gln276fs
M4	CDS: ecsB	c.1054_1055insA	p.Leu352fs
M5	CDS: rseP	c.1021C>T	p.Gln341*
M7	CDS: ecsB	c.117A>G	
M8	CDS: ecsB	c.636_638delGGCinsA	p.Gln212fs
M11	CDS: ecsB	c.106delG	p.Gly36fs
M12	CDS: ecsB	c.343delG	p.Gly115fs
M16	CDS: ecsB	c.1054_1055insA	p.Leu352fs
M19	CDS: ecsB	c.106delG	p.Gly36fs
M22	CDS: ecsA	c.640G>A, c.538delA	p.Val214Ile, p.Lys180fs

c) *L. garvieae* 1546

Isolate	Overlapping annotations	Coding region change	Amino acid change
G11	CDS: ecsA	c.187G>T	p.Gly63*
G12	CDS: rseP	c.119G>A	p.Gly40Glu
G13	CDS: ecsB	c.139C>T	p.Gln47*
G14	CDS: ecsA	c.465C>A	p.Tyr155*
G15	CDS: ecsA	c.172C>T	p.Gln58*
G16	CDS: ecsA	c.363C>A	p.Phe121Leu
G17	CDS: rseP	c.704G>A	p.Trp235*
G18	CDS: ecsB	c.380G>A	p.Trp127*
G19	CDS: ecsA	c.488G>T	p.Gly163Val
G20	CDS: rseP	c.448G>T	p.Glu150*

4.6 Assessment of resistance specificity

Due to the lack of mutations in the putative receptor gene *rseP*, some mutants were evaluated against the unrelated bacteriocins BHT-B and Garvicin ML to investigate if any unspecific resistance mechanism could have resulted from the non-*rseP*-mutations. These isolates were initially chosen from *E. faecalis* 3358 and *L. garvieae* 1546. The results suggested that no such general/unspecific mechanism was responsible for the resistance observed in the mutants without *rseP* alterations. The pattern of two different inhibition levels observed when using Ej97 was not present when exposing the isolates to the unrelated peptides instead (table 4.8).

Table 4.8: MIC_{50}-values attained for *E. faecalis* 3358 isolates when exposing them to Ej97 and BHT-B in the unspecific resistance assay. The isolates marked with * contain *rseP*-mutations. As the initial tests did not yield any significant differences, the experiment was not carried out for more isolates than those listed here.

Isolate:	Ej97 (µg/ml)	BHT-B (µg/ml)
3358 wt	0.8	0.8
E3*	>200	0.8
E25*	>200	0.4
E2	25	0.4
E4	25	0.4
E8	25	0.4
E11	25	0.4
E12	25	0.4

However, a few isolates from *E. faecalis* 3358 and *L. garvieae* 1546 were screened in the same manner using an active fraction of circular bacteriocin Garvicin ML with unknown concentration (data not shown). No differences in resistance were observed to this bacteriocin either between isolates with known *rseP* mutation, *ecsAB* mutation, or wild type. Taken together, these results indicated that the mutations caused resistance specific to only the leaderless bacteriocins.

4.7 Mutation frequency (*E. faecalis* 3358)

Mutation frequency analysis using mutation specific primers had been initiated and could perhaps be completed at a later point. 150 new *E. faecalis* 3358 mutants were produced for the experiment. The primers were designed to differentiate between one, two or three "CAAAAAAT" repeats, but were unfortunately not successful after several trials using template with known repeats, and it was decided to rather focus on other tasks.

5. DISCUSSION

5.1 Inhibition spectra

Implementing additional representatives of the LsbB-like motif in this assay could have been relevant in this experiment to gain additional information on activity-differences within this group of related peptides believed to interact with the same target receptor. Still, with the use of the three bacteriocins LsbB, K1 and Ej97, the results clearly demonstrated the contrasting activity of LsbB as a narrow-spectrum bacteriocin and Ej97 and K1 as broad-spectrum bacteriocins. While LsbB as expected only inhibited the growth of certain *Lactococcus*-strains, Ej97 and K1 exerted activity towards strains from various genera including *Listeria, Staphylococcus, Lactobacillus, Pediococcus, Lactococcus, Enterococcus.* Even between Ej97 and K1, subtle variations were present with Ej97 having the most comprehensive activity of the peptides. Ej97 demonstrated activity to certain strains K1 did not have any effect on, as well as frequently causing larger inhibition zones than K1.

Although this pattern of broad and narrow activity was more or less expected, these results are nevertheless interesting in demonstrating how peptides presumably targeting the same receptor molecule are distinguished by considerable differences in activity, which present knowledge cannot justify.

It should be noted that the results presented in table 4.1 report several suspect results, suggesting that errors, contaminations or mislabeling have been present either in the execution of the experiment, or in the LMG glycerol stock. Nisin should not be active on its own to Gram-negative species, implying that the results attained using *E. coli* LMG 2746 and *E. coli* LMG 3235 as indicators are misleading and faulty. The latter strain was also inhibited by the *Lactococcus*-specific LsbB which substantiates that unknown errors have affected the result. These indicator cultures in fact likely originate from totally different species. As two such errors were clearly visible and easily identified due to nisin implemented as a control, the possibility of additional but less obvious mistakes cannot be excluded. Preferably, the assay should have been repeated with elevated focus on preventing contaminations and perhaps sequencing/identification of indicator species in order to define the activity spectra with credible results.

5.2 The RseP receptor

While RseP is known to be a RIP (regulated intramembrane proteolysis) zinc-dependent metallopeptidase, the function and substrate of RseP is presently not established. In section 1.6, it was elaborated how the orthologous RseP in Gram-negative *E. coli* is well documented as involved in cell envelope stress response by cleavage of anti-sigma factor RseA facilitating release of σ^E, which initiates transcription of stress responses genes. One substrate of RseP in this species is hence RseA (Akiyama et al. 2004). In the Gram-positive *Bacillus subtilis*, the RseP homologue RasP is known to cleave anti-sigma factor RsiW to release σ^W in another stress response mediated mechanism (Heinrich et al. 2008). However, homology searches for both RseA and RsiW reveals that these genes are not present in the indicator species included in this study, making it apparent that RseP must conduct other tasks involving other substrates in these species.

Studies have confirmed that the lactococcal RseP (YvjB) functions as receptor for LsbB. In this study, it was concluded that RseP from a number of other species also act as receptor for Ej97, a relative of LsbB. RseP exists at homologues across many bacterial species, all containing more or less subtle differences. Many of these RseP proteins from the strains implemented in the inhibition spectrum assay, including the lactococcal RseP, seem to be recognized by broad-spectrum LsbB-like bacteriocins for target binding. In contrast, narrow-spectrum LsbB exclusively target the lactococcal RseP (figure 5.1). Hence, every LsbB-like bacteriocin seem to target an individually defined set of RseP proteins.

These activity differences are likely connected to differences in both the peptides and differences in the structure of the various RseP homologues. For target recognition, the bacteriocin structure needs to be sufficiently compatible with the receptor structure for the specific interaction to occur. In the case of LsbB, something in the LsbB structure seems to make in exclusively compatible with the lactococcal RseP. This is likely due to a distinct feature of the lactococcal receptor which is not present in other RseP proteins, and not necessary for interaction with Ej97 and K1.

Seeing as previous studies have indicated that truncated peptides can block the activity of LsbB as long as they contain the LsbB-like motif (KxxxGxxPWE), it seems that this motif is by far

most important for receptor binding. However, this motif is also present in all the investigated bacteriocins. Thus, characteristics in the peptides outside the conserved motif must somehow define the activity spectra by excluding compatibility with certain RseP proteins.

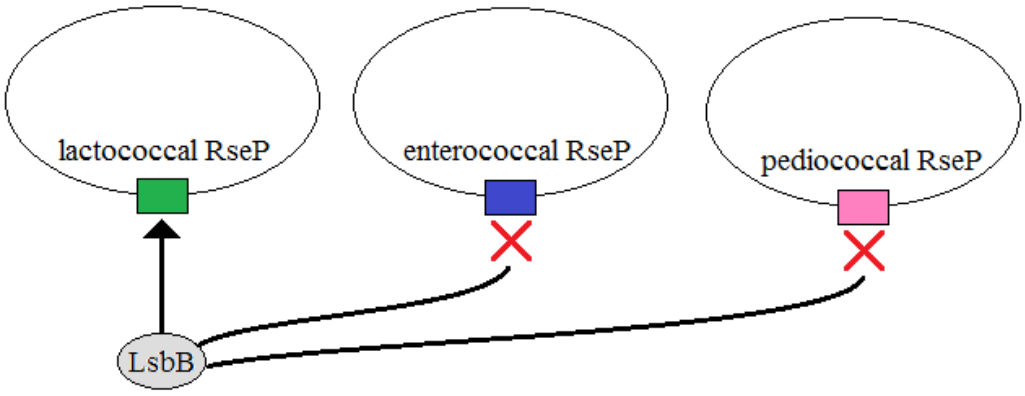
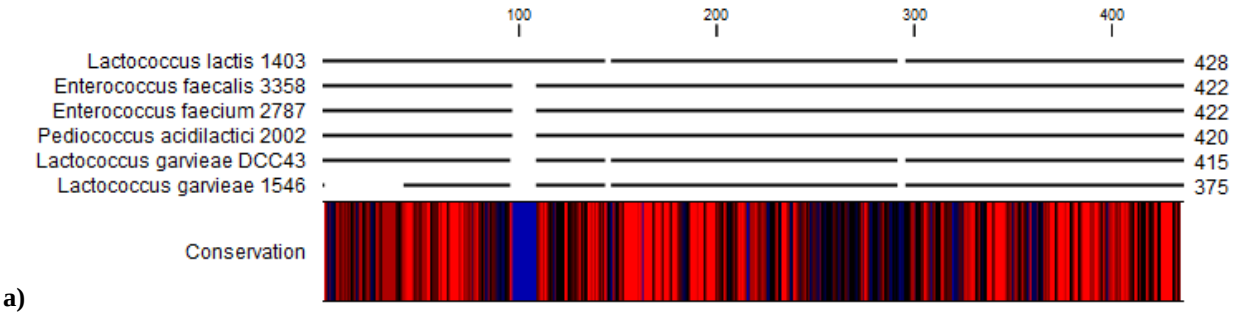


Figure 5.1: While Ej97 and K1 seem to have the ability to target RseP from a wide variety of bacterial species, LsbB is exclusively compatible with the lactococcal version (YvjB).

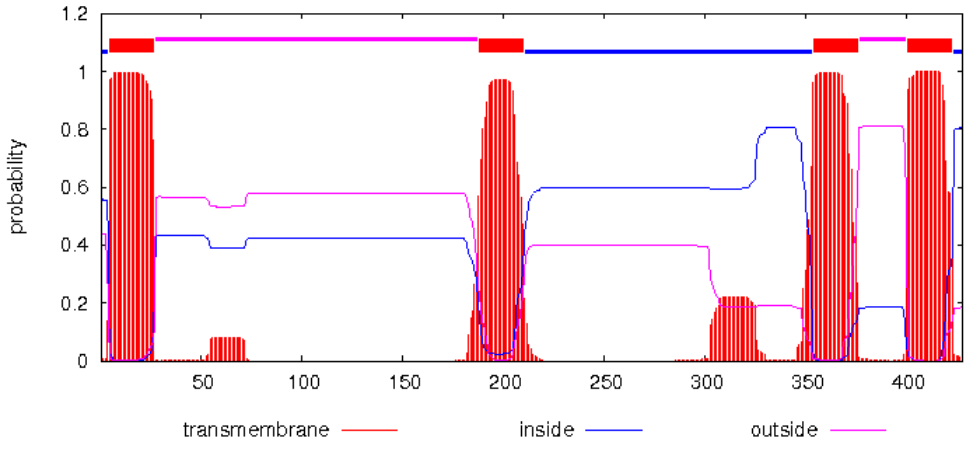
Comparison of YvjB and its RseP homologs from the indicator strains used throughout this study was carried out by making a multiple sequence alignment (MSA) of all the RseP proteins (figure 5.2a). The conservation graph of the MSA suggest that the LsbB-receptor YvjB has a sequence stretch ~100 aa into the sequence that is not present in the other RseP proteins. This stretch of ~13 aa (figure 5.2c) may be essential for LsbB-binding due to specific characteristics in the LsbB structure, while not necessary for binding with Ej97 and K1.

After predicting the transmembrane areas of the lactococcal RseP (figure 5.2b), the non-conserved sequence stretch appear to be located on the exterior of the membrane, thus in line with our notion that this region could be directly involved in receptor recognition. However, a more conclusive answer to this requires further investigation.

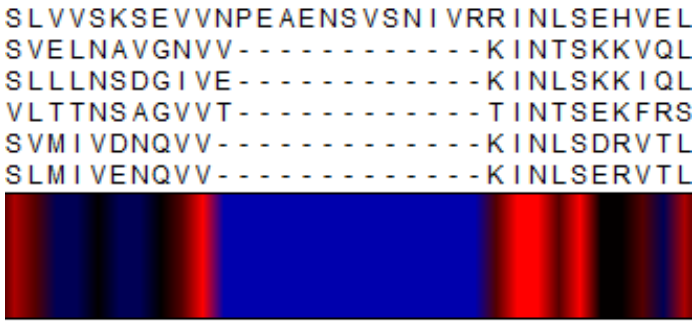
The figures presented in this section merely provide speculations and should be viewed critically. They only represent predictions based on the algorithms implemented in the various programs and are not based on factual evidence.



a)



b)



c)

Figure 5.2: a) The top picture shows how the lactococcal RseP (YvjB) in *L. lactis* IL1403 has a region approximately 100 aa into the protein which is not present/conserved across the other species implemented in this thesis. The picture is from CLC workbench after generating a MSA of the RseP proteins and displaying the degree of conservation as red (conserved) and blue (not conserved). b) The TMHMM prediction of the lactococcal RseP from *L. lactis* IL1403 displays how the protein is predicted to be oriented in the membrane according to the TMHMM algorithm. It also shows that the non-conserved sequence-stretch is predicted to be located at the outside of the membrane, which is where the initial bacteriocin interaction will occur. c) The bottom picture is a closer view of the blue area in a) and display the amino acid stretch which is apparently only present in the lactococcal receptor, and might somehow be needed for recognition by LsbB.

5.3 The phenotypes of naturally resistant mutants

The secondary task of the inhibition spectrum assay was to document indicator strains with resistant mutants occurring naturally when exposed to bacteriocin, and this determined which indicators to proceed with. Naturally resistant mutants from the two strains *E. faecalis* 3358 (Ej97-resistant) *and P. acidilactici* 2002 (Ej97- or K1-resistant) arose quite easily in response to exposure by Ej97/K1, and additional mutants from the three other strains *E. faecium* 2787, *L. garvieae* DCC43 and *L. garvieae* 1546 were supplemented later. It could have been exceedingly interesting to have put more effort into generating mutants from additional strains, and thus acquire more information on the resistant phenotypes versus genotypes. However, the consistency of the results from the five strains implemented here and their approximately 80 mutants provided sufficient proof of RseP as receptor for Ej97 and EcsAB as involved in medium Ej97-resistance.

Another aspect which would have been interesting to investigate, is the potential cross-resistance of the collected mutants towards other representatives of the LsbB-like motif. This was briefly done for certain isolates using K1 in microtiter plate assays, consistently suggesting presence of cross-resistance (results not shown). However, this was also indirectly investigated due to the fact that many of the mutants were collected from the inhibition zones of K1 or EntQ (table 4.2), and thus initially were K1- or EntQ-resistant rather than Ej97-resistant. As the microtiter plate results show, all mutants were still very much resistant to Ej97 despite this.

The most complete phenotype information was attained from investigating *E. faecalis* 3358 and *P. acidilactici* 2002. In these strains, the phenotypes clearly diverged in into two levels of resistance, which additionally corresponded perfectly with *rseP*-mutations (high resistance) and *ecsAB*-mutations (medium resistance). This suggested the presence of two different resistance mechanisms with different efficacies. This phenotype-genotype pattern could not be observed for the remaining three strains. However, this was not due to absence of the pattern and disproving the presence of such a connection, but merely a result of these mutants being resistant to higher bacteriocin concentrations. Thus, the microtiter assay in these instances could not properly conclude whether or not the mutants grouped into two resistance levels. Corresponding assays with higher concentrations could elucidate this. As a consequence, the phenotype results for

these three strains appear somewhat incomplete, and more effort should have been put into pursuing the determination of correct resistance levels (MIC_{50}-values).

The phenotype data from the microtiter assays were additionally not always consistent, making the results presented in table 4.3 somewhat speculative. With non-consistent results it is difficult to establish which result are faulty and not. The assays should preferably have been executed one more time to reassure that the MIC_{50}-values are correct and not a result of unknown errors.

5.4 *rseP* genotypes

In contrast to the results from the phenotype assays, which contain several shortcomings and need for improvements, the genotype results from sequencing were consistently pleasing and produced credible results in abundance.

Several isolates which had a very high resistance level compared to the wild type proved to also have serious mutations within the *rseP*-gene. The *rseP* mutations, although present in all the five strains, were very different in nature as described in the results section. However, the large majority of identified mutations resulted in premature termination of the protein. They ranged from simple point mutations causing missense, nonsense and frameshift mutations, to up-stream gene mutations, repeat variance, and transposons.

In the case of *L. garvieae* DCC43, all sequenced isolates contained nonsense or frameshift mutation within *rseP*, and whole genome sequencing was thus not carried out for this strain. Concerning *L. garvieae* 1546, comparison of the missense mutations in relation to phenotype made it apparent that Ej97 tolerates certain amino acid substitutions in the receptor, while other substitution confer resistance at the same level as premature termination. For instance, the *rseP* missense mutation in isolate G16 does not seem to cause a serious change in the receptor as the isolate maintains quite high sensitivity, while the missense mutation in G19 is as destructive to Ej97-sensitivity as premature termination. The same is evident for the missense mutations in M2 and M15 in *P. acidilactici* 2002 which also fail to confer a high level of resistance, while M10 with an additional missense mutation is as resistant as the mutants with truncated receptors. However, the majority of mutants completely lacked RseP mutations, suggesting that there had to be another resistance mechanism present in these isolates. With the two different resistance levels, there was likely also two different mechanisms of resistance.

5.5 Whole genome sequencing results

The WGS results suggested involvement of components from the Ecs ABC transporter in the medium resistant phenotype. This involvement was not confirmed experimentally in this thesis, but as *ecsAB* mutations were present in almost all medium resistance isolates and across species, the results are highly credible and should be pursued. *ecsAB*-mutations within 22 of the 23 sequenced medium resistance mutants (not including *E. faecium* 2787) can not be coincidental. The significance of these results are also reinforced by these genes having been established as interrelated in other works, as will be elaborated in the next section.

Mutations being present in *rseP* while not in *ecsAB* results in complete resistance to Ej97, but this is not true the other way around. Thus, the presence of functional EcsAB does not "help" the bacteriocin by providing a secondary receptor. If the RseP is mutated the bacterium is consistently highly resistant. When *ecsAB* is mutated and not *rseP*, the bacterium only attain medium resistance to Ej97.

It is likely that a mutation in *ecsAB* somehow influences the receptor RseP, either on transcription level (e.g. gene regulation), or on protein level (e.g. protein-protein interaction). A stress-response study on *B. subtilis* and its RseP orthologue RasP argue that it is not on transcription level that the *ecsAB* interferes with *rseP*, but rather that the effect is exerted on protein level and interfering with mature RasP. This will be clarified in the next section.

In general, the whole genome sequencing was very successful with high coverage assisting good assemblies. However, the *E. faecium* 2787 wild type attained somewhat poorer coverage and after several assembly attempts it was decided to not put more effort into improving the assembly of the wild type. This was very unfortunate as the wild types are in greatest need of good coverage as they are the basis all the mutants will be evaluated against. For another time, the WGS sample preparation should make sure the wild types will receive a higher coverage than the other samples, to guarantee that the data are sufficient for good assembly. Luckily, the three other strains subjected to WGS provided good evidence that *ecsAB* in some manner must be involved in causing resistance to Ej97. Also, *ecsA* mutations were identified in or up-stream 3 of the 8 sequenced *E. faecium* 2787 mutants by searching specifically in this gene (after making the discovery in the other strains), and this is likely also not coincidental.

5.6 Involvement of EcsAB in resistance

The main function of ABC transporters is to translocate substances in or out of cells and organelles. They belong to one of the largest protein superfamilies and are present in multiple variants in all organisms. ABC transporters consist of several subunits, including hydrophobic components spanning the membrane, and hydrophilic components facilitation ATP hydrolysis to couple the energy yield with the transportation mechanism.

EcsAB is such an ABC transporter, but with elusive function and facilitating transport of unknown substrates. TMHMM predictions (data not shown) suggest that *ecsA* is not affiliated with the membrane whilst *ecsB* has several membrane-spanning domains. After WGS, mutations were discovered both within the A and B component of *ecsAB,* where EcsB is the transmembrane permease protein. EcsA is the hydrophilic ATP-binding component involved in providing energy for the transport-function of the B-component.

The results of this study indicated a correlation between EcsAB and bacteriocin resistance, more specifically medium resistance, likely by EcsAB-mutations interfering with the receptor RseP. A study from 2008 on *B. subtilis* and stress response contains interesting information in regards this discovery (Heinrich et al. 2008):

Genes within an antibiosis regulon in *B. subtilis* responding to a variety of stress signals on the cellular envelope are induced by an alternative sigma factor σ^W. The activity of σ^W is controlled by RsiW which is a transmembrane anti-sigma factor capable of inactivating σ^W. If a stress signal is encountered, the RsiW is degraded by RasP (zinc-dependent intermembrane proteolysis) and two other proteases, in order to release σ^W into the cell where it will subsequently bind with RNA polymerase at the promotor site of the stress regulon. In this manner it induces transcription of the antibiosis regulon and help the cell react to the stress signal. RasP is an orthologue of RseP which is present in many bacterial species.

The results of the *B. subtilis*-study indicated that mutations within the EcsAB transporter block the cleavage of anti-sigma factor RsiW by somehow interfering with the intramembrane protease RasP responsible for the cleavage. This means that when EcsAB is mutated, the bacterium is unable to cleave RsiW which will continue to repress σ^W, and in turn inhibit transcription of the genes needed for response to the stress signal.

How EcsAB interferes with RasP is not known. The *ecsAB* mutations in *B. subtilis* were suggested to cause a defect in the RasP function by the protein not being properly inserted in the membrane and no longer carrying out its task of cleaving RsiW properly. It was also suggested that EcsAB as an ABC transporter may be involved in translocation and inserting other components to the cell membrane, which when not done properly affects RasP in the membrane as well. It was additionally proposed that it could be the substrate of EcsAB that interferes with RasP when the EcsAB in non-functional, and hence inhibits functionality of RasP as well. The results from the study favor a model where the absence of functional EcsAB inhibits existing RasP from carrying out is function, and not a model where EcsAB is linked to having a regulatory effect over expression levels of RasP (Heinrich et al. 2008).

In another study, focusing on *S. aureus* strains with mutated *ecsAB,* it was discovered that mutations in *ecsAB* caused growth defects and altered cell wall composition and surface texture, as well as increased autolysis and lysostaphin sensitivity. Additionally, the mutants had changed expression of the virulence factor Rot, and were observed to cause markedly milder infections in mice. The *ecsAB* mutants were at the same time more susceptible to ribosomal antibiotics. The researchers concluded that the ecs operon is essential for the normal structure and function of the cell wall. (Jonsson et al. 2010). This further substantiates the notion that EcsAB-deficiency can affect mature RseP, and in relation to this present study somehow makes it inaccessible to Ej97.

As mentioned in the introduction, altering the composition and structure of the cell membrane may render a bacterium resistant to bacteriocin by restricting the peptide from reaching its target, and this could possibly have relevance in context to the EcsAB-effect on intact RseP. Alternatively, the EcsAB transporter could be required for the removal of the degraded products resulting from RseP proteolysis, and the removal of RseP-mediated products being important for regeneration of RseP function. For functionality as receptor for Ej97, the RseP protease needs to be intact without mutations altering the extracellular binding site of Ej97, and this binding site must also be available for the bacteriocin to dock with it, i.e. not bound by other substrates, oriented differently in the membrane, or otherwise protected from the bacteriocin by modifications in the membrane structure. However, these are all merely speculations at the present time. As will be elaborated in the next section, the resistance caused *ecsAB*-mutations also seemed to be specific to Ej97.

5.7 Mutations identified give rise to specific resistance

While the results from sequencing supported the presence of two different resistance mechanisms causing two different phenotypes, it was not yet clear if both these resistance mechanisms specifically inhibited the action of Ej97. There was also a possibility of unspecific mechanisms, which include mechanisms involving increased protease activity for degrading extracellular substances or efflux pump activity. As the Ecs ABC transporter potentially hold an efflux function for exportation, this would be the most likely scenario. However, the EcsAB was in this case mutated, which would suggest that its activity was impaired, not increased.

If unspecific mechanisms were responsible for the medium resistance, one could expect the medium resistant isolates to have increased resistance to other bacteriocins not targeting the same receptor as well.

However, the few results obtained in this assay regarding this all indicated that this was not the case, and that the medium resistance also was related specifically to avoiding the action of Ej97. No increase in resistance was observed within the medium resistant isolates (or the wild type and RseP mutants) when exposing them to other bacteriocins. With Ej97, a specific patter of different resistance levels had been observed, but this was not present when using the other bacteriocin BHT-B and Garvicin ML. Thus, the results indicate that both the *rseP*- and *ecsAB*-mutations confer Ej97-specific resistance mechanisms.

Still, this experiment is not sufficient to draw this definitive conclusion. More unrelated bacteriocins as well as more resistant isolates should have been employed in the assay, in order to be certain about this matter. Once again, time was the limiting factor and this particular experiment was not of high priority for the time being but could perhaps be expanded and finished at a later point.

5.8 A summary of genotype and phenotype results

Below, in table 5.1 is a summary the results collected for this thesis presented. The isolates are organized according to resistance level (descending) to better signify how RseP-mutations are related to high resistance whilst EcsAB-mutations confer medium resistance. This relationship is as discussed earlier most evident in *E. faecalis* 3358 and *P. acidilactici* 2002. This is a consequence of the remaining strains having less complete phenotype results and thus such a connection cannot be made based on these results.

Table 5.1: Overview of mutations within the *rseP* and *ecsAB* in the sequenced isolates, arranged according to their resistance levels. The table displays how the isolates with receptor gene *rseP*-mutations generally have highest resistance to Ej97, while *ecsAB*-mutation cause a lower level of resistance (but significantly higher than in the wild types). "-"denotes isolates where no mutations were identified in the respective genes.

Indicator bacteria:	Isolate:	CFU picked from inhibition zone-type:	Ej97 MIC_{50} value (µg/ml)	*rseP*-mutations	*ecsAB*-mutations (WGS)
E. faecalis 3358	E3	Ej97 0,25	>50	Frameshift	-
	E10	Ej97 0,5	>50	Frameshift	*(not sequenced)*
	E17	Ej97 0,1	>50	Frameshift	*(not sequenced)*
	E20	Ej97 0,5	>50	Frameshift	*(not sequenced)*
	E22	Ej97 0,5	>50	Frameshift	*(not sequenced)*
	E25	Ej97 0,25	>50	Frameshift	-
	E28	Ej97 0,1	>50	Frameshift	*(not sequenced)*
	E4	Ej97 0,25	50	-	-
	E8	Ej97 0,1	50	-	Frameshift *ecsA*
	E14	Ej97 0,25	50	-	Frameshift *ecsA*
	E15	Ej97 0,1	50	-	*(not sequenced)*
	E7	Ej97 0,25	25	-	*(not sequenced)*
	E12	Ej97 0,25	25	-	Frameshift *ecsA*
	E23	Ej97 0,25	25	-	*(not sequenced)*
	E24	Ej97 0,25	25	-	Frameshift *ecsA*
	E26	Ej97 0,1	25	-	Frameshift *ecsA*
	E27	Ej97 0,1	25	-	*(not sequenced)*
	E1	Ej97 0,5	12.5	-	*(not sequenced)*
	E2	Ej97 0,5	12.5	-	Nonsense *ecsA*
	E11	Ej97 0,5	12.5	-	Frameshift *ecsA*
	3358 WT		0.4		
P. acidilactici 2002	M1	Ej97 0,5	>50	Premature termination	-
	M5	Ej97 0,25	>50	Premature termination	-
	M6	Ej97 0,25	>50	Premature termination	*(not sequenced)*
	M17	Ej97 0,5	>50	Upstream gene mutation	*(not sequenced)*
	M18	Ej97 0,25	>50	Premature termination	*(not sequenced)*
	M20	Ej97 0,25	>50	Premature termination	*(not sequenced)*
	M21	Ej97 0,25	>50	Premature termination	*(not sequenced)*
	M10	K1 0,5	25	Amino acid substitution	*(not sequenced)*
	M15	Ej97 0,5	12.5	Amino acid substitution	*(not sequenced)*
	M7	Ej97 0,1	12.5	-	Frameshift *ecsB*

	M8	Ej97 0,1	12.5	-	Frameshift *ecsB*
	M11	K1 0,5	12.5	-	Frameshift *ecsB*
	M12	K1 0,5	6,25	-	Frameshift *ecsB*
	M2	Ej97 0,5	6.25	Amino acid substitution	*(not sequenced)*
	M9	Ej97 0,1	6.25	-	*(not sequenced)*
	M22	Ej97 0,25	6.25	-	Frameshift *ecsA*
	M3	Ej97 0,5	3.125	-	*(not sequenced)*
	M4	Ej97 0,25	3.125	-	Frameshift *ecsB*
	M16	Ej97 0,5	3.125	-	Frameshift *ecsB*
	M19	Ej97 0,25	3.125	-	*(not sequenced)*
	M13	Ej97 0,5	3.125	-	*(not sequenced)*
	M14	Ej97 0,5	3.125	-	*(not sequenced)*
	2002 WT		0.2		
L. garvieae DCC43	G1	Ej97	>50	Frameshift type I	*(not sequenced)*
	G2	Ej97	>50	Nonsense mutation	*(not sequenced)*
	G3	Ej97	>50	Frameshift type II	*(not sequenced)*
	G4	Ej97	>50	Frameshift type I	*(not sequenced)*
	G5	Ej97	>50	Nonsense mutation	*(not sequenced)*
	G6	K1	>50	Nonsense mutation	*(not sequenced)*
	G7	K1	>50	Nonsense mutation	*(not sequenced)*
	G8	K1	>50	Frameshift type II	*(not sequenced)*
	G9	K1	>50	Frameshift type II	*(not sequenced)*
	G10	K1	>50	Frameshift type I	*(not sequenced)*
	DCC43 WT		0.8		
L. garvieae 1546	G11	Ej97	>50	-	Nonsense *ecsA*
	G12	Ej97	>50	Upstream gene mutation	-
	G13	Ej97	>50	-	Nonsense *ecsB*
	G14	Ej97	>50	-	Nonsense *ecsA*
	G15	Ej97	>50	-	Nonsense *ecsA*
	G17	K1	>50	Nonsense mutation	-
	G18	K1	>50	-	Nonsense *ecsB*
	G19	K1	>50	-	Missense *ecsA*
	G20	K1	>50	Nonsense mutation	-
	G16	K1	6.25	-	Missense *ecsA*
	1546 WT		6.25		
E. faecium 2787	B30	EntQ	>50	Transposon	*(not sequenced)*
	B32	EntQ	>50	Transposon	*(not sequenced)*
	B70	K1	>50	Transposon	*(not sequenced)*
	B72	K1	>50	-	*ecsA*mutation
	B74	K1	>50	Transposon	*(not sequenced)*
	B76	K1	>50	Transposon	*(not sequenced)*
	B80	K1	>50	Transposon	*ecsA*mutation
	B84	K1	>50	Transposon	-
	B94	K1	>50	Transposon	*(not sequenced)*
	B86	K1	>50	-	-
	B36	EntQ	>50	-	-
	B44	EntQ	>50	-	-
	B28	EntQ	>50	-	-
	B38	EntQ	>50	-	Upstream *ecsA*mutation
	B42	EntQ	>50	-	*(not sequenced)*
	B48	EntQ	>50	-	-
	2787 WT		0.6		

5.9 Concluding remarks and future prospects

This thesis presents the successful identification of the RIP metallopeptidase RseP as receptor for the LsbB-like bacteriocin Ej97. This conclusion was made by observing that high Ej97-resistance was linked with severe RseP-mutations in five bacterial strains from four different species. This was a somewhat expected result due to the distinct similarities between Ej97 and its *Lactococcus*-specific relative LsbB. However, there was also observed a medium resistance level for which *rseP*-mutations were not accountable.

Additionally, the results present interesting new leads in target recognition for leaderless bacteriocins by providing clear indications of a second resistance mechanism. After completing WGS, mutations within the Ecs ABC transporter of at least three different species seem to be linked directly to medium resistance to Ej97. This proposes a previously unknown mechanism, one which is likely directed specifically towards the bacteriocin but at the same time less effective than the resistance achieved by receptor gene mutation.

A previous study on a system not related to bacteriocins have unraveled a relationship between EcsAB and an RseP-homologue. This study describes mutations in EcsAB rendering the RseP less active and unable to carry out it designated task in cleaving substrates. A similar relationship could potentially be present in the strains investigated here, and ultimately conferring medium bacteriocin resistance. Exactly how EcsAB interferes with RseP is presently unknown (figure 5.3) but should be further investigated (Heinrich et al. 2008).

What remains to be elucidated are the roles of these two protein and how *ecsAB* mutations affect *rseP*, making them interrelated in bacteriocin resistance. Further prospects include investigating if *ecsAB* mutations affect the RseP on transcription level or protein level in some manner. Some possibilities involve qPCR to evaluate the expression level of the RseP receptor when *ecsAB* is mutated (the possibility of less RseP proteins being produced as a consequences of the *ecsAB* being mutated), or diverse cloning experiments to investigate the relationship between EcsAB and mature RseP.

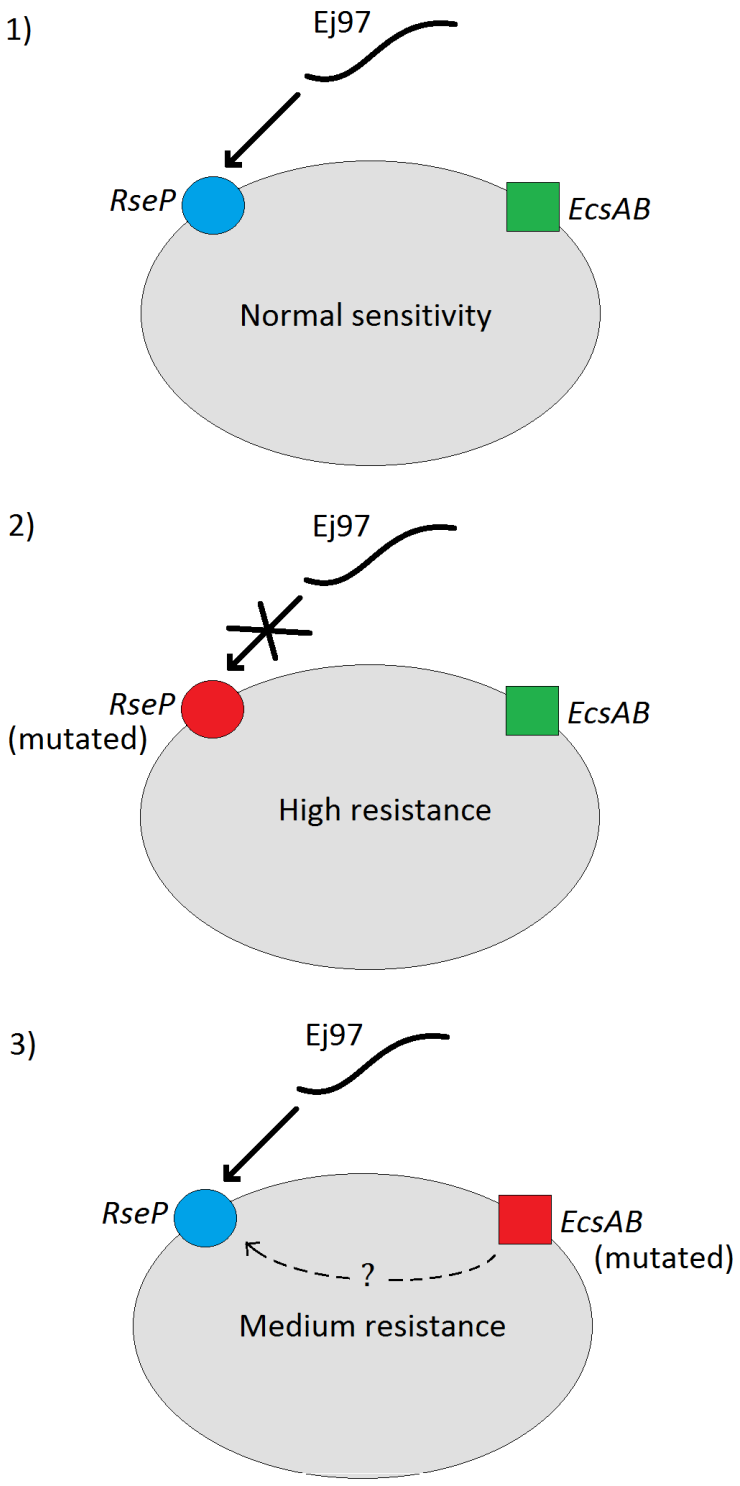


Figure 5.3: A simple illustration for Ej97-resistance. When RseP is mutated the bacterium attain high level of resistance to the Ej97, while when the Ecs ABC transporter is mutated whilst the RseP intact, the bacterium somehow attain medium resistance.

6. REFERENCES

Akiyama, Y., Kanehara, K. & Ito, K. (2004). RseP (YaeL), an Escherichia coli RIP protease, cleaves transmembrane sequences. *EMBO J*, 23 (22): 4434-42.

APUA, A. f. t. P. U. o. A. (2014). *General Background: Antibiotic Agents*. Available at: http://www.tufts.edu/med/apua/about_issue/agents.shtml (accessed: 16.11.14).

Borrero, J., Brede, D. A., Skaugen, M., Diep, D. B., Herranz, C., Nes, I. F., Cintas, L. M. & Hernandez, P. E. (2011). Characterization of garvicin ML, a novel circular bacteriocin produced by Lactococcus garvieae DCC43, isolated from mallard ducks (Anas platyrhynchos). *Appl Environ Microbiol*, 77 (1): 369-73.

Cheng, G., Hao, H., Xie, S., Wang, X., Dai, M., Huang, L. & Yuan, Z. (2014). Antibiotic alternatives: the substitution of antibiotics in animal husbandry? *Front Microbiol*, 5: 217.

Cintas, L. M., Casaus, P., Herranz, C., Havarstein, L. S., Holo, H., Hernandez, P. E. & Nes, I. F. (2000). Biochemical and genetic evidence that Enterococcus faecium L50 produces enterocins L50A and L50B, the sec-dependent enterocin P, and a novel bacteriocin secreted without an N-terminal extension termed enterocin Q. *J Bacteriol*, 182 (23): 6806-14.

Cotter, P. D., Ross, R. P. & Hill, C. (2013). Bacteriocins - a viable alternative to antibiotics? *Nat Rev Microbiol*, 11 (2): 95-105.

Cotter, P. D. (2014). An 'Upp'-turn in bacteriocin receptor identification. *Mol Microbiol*, 92 (6): 1159-63.

Diep, D. B., Axelsson, L., Grefsli, C. & Nes, I. F. (2000). The synthesis of the bacteriocin sakacin A is a temperature-sensitive process regulated by a pheromone peptide through a three-component regulatory system. *Microbiology*, 146 (Pt 9): 2155-60.

Diep, D. B., Skaugen, M., Salehian, Z., Holo, H. & Nes, I. F. (2007). Common mechanisms of target cell recognition and immunity for class II bacteriocins. *Proc Natl Acad Sci U S A*, 104 (7): 2384-9.

Draper, L. A., Ross, R. P., Hill, C. & Cotter, P. D. (2008). Lantibiotic immunity. *Curr Protein Pept Sci*, 9 (1): 39-49.

Drider, D., Fimland, G., Hechard, Y., McMullen, L. M. & Prevost, H. (2006). The continuing story of class IIa bacteriocins. *Microbiol Mol Biol Rev*, 70 (2): 564-82.

Duquesne, S., Destoumieux-Garzon, D., Peduzzi, J. & Rebuffat, S. (2007). Microcins, gene-encoded antibacterial peptides from enterobacteria. *Nat Prod Rep*, 24 (4): 708-34.

Eijsink, V. G., Axelsson, L., Diep, D. B., Havarstein, L. S., Holo, H. & Nes, I. F. (2002). Production of class II bacteriocins by lactic acid bacteria; an example of biological warfare and communication. *Antonie Van Leeuwenhoek*, 81 (1-4): 639-54.

Fimland, G., Johnsen, L., Dalhus, B. & Nissen-Meyer, J. (2005). Pediocin-like antimicrobial peptides (class IIa bacteriocins) and their immunity proteins: biosynthesis, structure, and mode of action. *J Pept Sci*, 11 (11): 688-96.

Gabrielsen, C., Brede, D. A., Hernandez, P. E., Nes, I. F. & Diep, D. B. (2012). The maltose ABC transporter in Lactococcus lactis facilitates high-level sensitivity to the circular bacteriocin garvicin ML. *Antimicrob Agents Chemother*, 56 (6): 2908-15.

Gabrielsen, C., Brede, D. A., Nes, I. F. & Diep, D. B. (2014). Circular bacteriocins: biosynthesis and mode of action. *Appl Environ Microbiol*, 80 (22): 6854-62.

Gajic, O., Buist, G., Kojic, M., Topisirovic, L., Kuipers, O. P. & Kok, J. (2003). Novel mechanism of bacteriocin secretion and immunity carried out by lactococcal multidrug resistance proteins. *J Biol Chem*, 278 (36): 34291-8.

Galvez, A., Valdivia, E., Abriouel, H., Camafeita, E., Mendez, E., Martinez-Bueno, M. & Maqueda, M. (1998). Isolation and characterization of enterocin EJ97, a bacteriocin produced by Enterococcus faecalis EJ97. *Arch Microbiol*, 171 (1): 59-65.

Galvez, A., Abriouel, H., Lopez, R. L. & Ben Omar, N. (2007). Bacteriocin-based strategies for food biopreservation. *Int J Food Microbiol*, 120 (1-2): 51-70.

Garneau, S., Martin, N. I. & Vederas, J. C. (2002). Two-peptide bacteriocins produced by lactic acid bacteria. *Biochimie*, 84 (5-6): 577-92.

Hassan, M., Kjos, M., Nes, I. F., Diep, D. B. & Lotfipour, F. (2012). Natural antimicrobial peptides from bacteria: characteristics and potential applications to fight against antibiotic resistance. *J Appl Microbiol*, 113 (4): 723-36.

Hechard, Y. & Sahl, H. G. (2002). Mode of action of modified and unmodified bacteriocins from Gram-positive bacteria. *Biochimie*, 84 (5-6): 545-57.

Heinrich, J., Lunden, T., Kontinen, V. P. & Wiegert, T. (2008). The Bacillus subtilis ABC transporter EcsAB influences intramembrane proteolysis through RasP. *Microbiology*, 154 (Pt 7): 1989-97.

Huddleston, J. R. (2014). Horizontal gene transfer in the human gastrointestinal tract: potential spread of antibiotic resistance genes. *Infect Drug Resist*, 7: 167-76.

Hyink, O., Balakrishnan, M. & Tagg, J. R. (2005). Streptococcus rattus strain BHT produces both a class I two-component lantibiotic and a class II bacteriocin. *FEMS Microbiol Lett*, 252 (2): 235-41.

Jack, R. W., Tagg, J. R. & Ray, B. (1995). Bacteriocins of gram-positive bacteria. *Microbiol Rev*, 59 (2): 171-200.

Kjos, M., Nes, I. F. & Diep, D. B. (2009). Class II one-peptide bacteriocins target a phylogenetically defined subgroup of mannose phosphotransferase systems on sensitive cells. *Microbiology*, 155 (Pt 9): 2949-61.

Kjos, M., Salehian, Z., Nes, I. F. & Diep, D. B. (2010). An extracellular loop of the mannose phosphotransferase system component IIC is responsible for specific targeting by class IIa bacteriocins. *J Bacteriol*, 192 (22): 5906-13.

Kjos, M., Oppegard, C., Diep, D. B., Nes, I. F., Veening, J. W., Nissen-Meyer, J. & Kristensen, T. (2014). Sensitivity to the two-peptide bacteriocin lactococcin G is dependent on UppP, an enzyme involved in cell-wall synthesis. *Mol Microbiol*, 92 (6): 1177-87.

Kroos, L. & Akiyama, Y. (2013). Biochemical and structural insights into intramembrane metalloprotease mechanisms. *Biochim Biophys Acta*, 1828 (12): 2873-85.

Lohans, C. T. & Vederas, J. C. (2012). Development of Class IIa Bacteriocins as Therapeutic Agents. *Int J Microbiol*, 2012: 386410.

Martin, N. I. & Breukink, E. (2007). Expanding role of lipid II as a target for lantibiotics. *Future Microbiol*, 2 (5): 513-25.

Nes, I. F., Diep, D. B., Havarstein, L. S., Brurberg, M. B., Eijsink, V. & Holo, H. (1996). Biosynthesis of bacteriocins in lactic acid bacteria. *Antonie Van Leeuwenhoek*, 70 (2-4): 113-28.

Nes, I. F., Yoon, S. & Diep, D. B. (2007). Ribosomally Synthesiszed Antimicrobial Peptides (Bacteriocins) in Lactic Acid Bacteria: A Review. *Food Science and Biotechnology*, 16 (5): 675-690.

Nissen-Meyer, J. & Nes, I. F. (1997). Ribosomally synthesized antimicrobial peptides: their function, structure, biogenesis, and mechanism of action. *Arch Microbiol*, 167 (2/3): 67-77.

Ovchinnikov, K. V., Kristiansen, P. E., Uzelac, G., Topisirovic, L., Kojic, M., Nissen-Meyer, J., Nes, I. F. & Diep, D. B. (2014). Defining the structure and receptor binding domain of the leaderless bacteriocin LsbB. *J Biol Chem*, 289 (34): 23838-45.

Reeves, P. (1965). The Bacteriocins. *Bacteriol Rev*, 29: 24-45.

Snyder, A. B. & Worobo, R. W. (2014). Chemical and genetic characterization of bacteriocins: antimicrobial peptides for food safety. *J Sci Food Agric*, 94 (1): 28-44.

Stevens, K. A., Sheldon, B. W., Klapes, N. A. & Klaenhammer, T. R. (1991). Nisin treatment for inactivation of Salmonella species and other gram-negative bacteria. *Appl Environ Microbiol*, 57 (12): 3613-5.

Tagg, J. R., Dajani, A. S. & Wannamaker, L. W. (1976). Bacteriocins of gram-positive bacteria. *Bacteriol Rev*, 40 (3): 722-56.

Todar, K. (2008). *Control of Microbial Growth* Available at: http://textbookofbacteriology.net/control_1.html (accessed: 09.12.14).

Uzelac, G., Kojic, M., Lozo, J., Aleksandrzak-Piekarczyk, T., Gabrielsen, C., Kristensen, T., Nes, I. F., Diep, D. B. & Topisirovic, L. (2013). A Zn-dependent metallopeptidase is responsible for sensitivity to LsbB, a class II leaderless bacteriocin of Lactococcus lactis subsp. lactis BGMN1-5. *J Bacteriol*, 195 (24): 5614-21.

Wiedemann, I., Breukink, E., van Kraaij, C., Kuipers, O. P., Bierbaum, G., de Kruijff, B. & Sahl, H. G. (2001). Specific binding of nisin to the peptidoglycan precursor lipid II combines pore formation and inhibition of cell wall biosynthesis for potent antibiotic activity. *J Biol Chem*, 276 (3): 1772-9.

Yazdankhah, S., Lassen, J., Midtvedt, T. & Solberg, C. O. (2013). [The history of antibiotics]. *Tidsskr Nor Laegeforen*, 133 (23-24): 2502-7.

APPENDIX

Attachments:

A1: *E. faecalis* 3358 *rseP* (MSA)

A2: *E. faecalis* 3358 RseP (MSA)

A3: *P. acidilactici* 2002 *rseP* (MSA)

A4: *P. acidilactici* 2002 RseP (MSA)

A5: *L. garvieae* DCC43 *rseP* (MSA)

A6: *L. garvieae* DCC43 RseP (MSA)

A7: *L. garvieae* 1546 *rseP* (MSA)

A8: *L. garvieae* 1546 RseP (MSA)

A9: *E. faecium* 2787 transposons (MSA)

A10: truncated mutants and TMHMM prediction of *E. faecalis* 3358 RseP

A11: truncated mutants and TMHMM prediction of *P. acidilactici* 2002 RseP

A12: truncated mutants and TMHMM prediction of *L. garvieae* DCC43 RseP

A13: truncated mutants and TMHMM prediction of *L. garvieae* 1546 RseP

A14: truncated mutants and TMHMM prediction of *E. faecium* 2787 RseP

A15: *E. faecalis* 3358 EcsA (MSA)

A16: *P. acidilactici* 2002 EcsA (MSA)

A17: *L. garvieae* 1546 EcsA (MSA)

A18: *P. acidilactici* 2002 EcsB (MSA)

A19: *E. faecalis* 3358 EcsB (MSA)

A20: *L. garvieae* 1546 EcsB (MSA)

A21: *E. faecalis* 3358 Probabilistic variant detection results

A22: *L. garvieae* 1546 Probabilistic variant detection results

A23: *P. acidilactici* 2002 Probabilistic variant detection results

Enterococcus faecalis 3358

A1: MSA of 3358 *rseP*

```
WT3358          ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E1              ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E2              ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E3              ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E4              ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E7              ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E8              ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E10             ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E11             ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E12             ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E14             ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E15             ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E17             ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E20             ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E22             ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E23             ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E24             ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E25             ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E26             ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E27             ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
E28             ATGAAGGAAGAAAGAAGGACATTTATGAAAACAATTATCACATTCATTATTGTCTTCGGT 60
                ************************************************************

WT3358          ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E1              ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E2              ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E3              ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E4              ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E7              ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E8              ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E10             ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E11             ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E12             ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E14             ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E15             ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E17             ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E20             ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E22             ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E23             ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E24             ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E25             ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E26             ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E27             ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
E28             ATTCTTGTCCTCGTACATGAATTTGGCCACTTTTATTTTGCAAAACGAGCGGGTATTTTA 120
                ************************************************************

WT3358          GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E1              GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E2              GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E3              GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E4              GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E7              GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E8              GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E10             GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E11             GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E12             GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E14             GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E15             GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E17             GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E20             GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E22             GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E23             GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E24             GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E25             GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E26             GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E27             GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
E28             GTTCGTGAGTTTGCAATCGGAATGGGACCAAAGATTTTTGCGCATCGTGGAAAAGATGGC 180
                ************************************************************

WT3358          ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E1              ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E2              ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E3              ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E4              ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E7              ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E8              ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E10             ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E11             ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E12             ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E14             ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E15             ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E17             ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E20             ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E22             ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E23             ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
```

```
E24             ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E25             ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E26             ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E27             ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
E28             ACCACTTATACGATTCGCTTATTGCCGATTGGTGGCTATGTGCGAATGGCTGGGATGGGC 240
                ************************************************************

WT3358          GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E1              GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E2              GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E3              GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E4              GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E7              GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E8              GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E10             GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E11             GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E12             GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E14             GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E15             GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E17             GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E20             GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E22             GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E23             GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E24             GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E25             GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E26             GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E27             GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
E28             GAAGACATGACAGAAATTACACCAGGTATGCCTCTATCTGTTGAGTTAAATGCCGTGGGT 300
                ************************************************************

WT3358          AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E1              AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E2              AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E3              AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E4              AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E7              AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E8              AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E10             AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E11             AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E12             AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E14             AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E15             AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E17             AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E20             AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E22             AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E23             AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E24             AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E25             AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E26             AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E27             AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
E28             AATGTGGTTAAAATTAATACAAGTAAAAAAGTACAATTACCTCATAGTATTCCGATGGAA 360
                ************************************************************

WT3358          GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E1              GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E2              GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E3              GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E4              GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E7              GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E8              GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E10             GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E11             GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E12             GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E14             GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E15             GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E17             GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E20             GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E22             GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E23             GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E24             GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E25             GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E26             GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E27             GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
E28             GTCGTTGATTTTGATCTTGAAAAAGAATTATTCATCAAGGGCTATGTCAATGGAAACGAG 420
                ************************************************************

WT3358          GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E1              GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E2              GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E3              GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E4              GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E7              GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E8              GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E10             GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E11             GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E12             GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E14             GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E15             GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E17             GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E20             GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E22             GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E23             GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
```

```
E24             GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E25             GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E26             GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E27             GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
E28             GAAGAAGAAACCGTTTATAAAGTTGACCATGATGCAACGATTATTGAAAGTGATGGTACC 480
                ************************************************************

WT3358          GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E1              GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E2              GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E3              GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E4              GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E7              GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E8              GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E10             GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E11             GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E12             GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E14             GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E15             GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E17             GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E20             GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E22             GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E23             GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E24             GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E25             GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E26             GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E27             GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
E28             GAGGTGCGGATTGCGCCACTTGACGTTCAATTTCAATCAGCGAAATTATCGCAACGCATT 540
                ************************************************************

WT3358          TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E1              TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E2              TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E3              TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E4              TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E7              TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E8              TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E10             TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E11             TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E12             TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E14             TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E15             TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E17             TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E20             TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E22             TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E23             TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E24             TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E25             TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E26             TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E27             TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
E28             TTAACGAACTTTGCGGGACCCATGAATAACTTTATCTTAGGGTTTATTCTGTTTACGTTA 600
                ************************************************************

WT3358          GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E1              GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E2              GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E3              GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E4              GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E7              GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E8              GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E10             GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E11             GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E12             GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E14             GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E15             GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E17             GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E20             GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E22             GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E23             GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E24             GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E25             GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E26             GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E27             GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
E28             GCGGTCTTTCTACAAGGAGGCGTCACTGATTTAAACACGAATCAAATTGGACAAGTGATT 660
                ************************************************************

WT3358          CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E1              CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E2              CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E3              CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E4              CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E7              CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E8              CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E10             CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E11             CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E12             CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E14             CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E15             CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E17             CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E20             CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E22             CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E23             CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
```

```
E24             CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E25             CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E26             CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E27             CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
E28             CCTAATGGCCCAGCCGCAGAAGCTGGGTTGAAAGAAAACGATAAAGTCTTATCGATTAAT 720
                ************************************************************

WT3358          AATCAAAAAATCAAAAAAT--------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 772
E1              AATCAAAAAATCAAAAAAT--------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 772
E2              AATCAAAAAATCAAAAAAT--------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 772
E3              AATCAAAAAAT----------------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 764
E4              AATCAAAAAATCAAAAAAT--------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 772
E7              AATCAAAAAATCAAAAAAT--------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 772
E8              AATCAAAAAATCAAAAAAT--------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 772
E10             AATCAAAAAATCAAAAAATCAAAAAATACGAAGATTTTACAACCATTGTGCAGAAGAACC 780
E11             AATCAAAAAATCAAAAAAT--------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 772
E12             AATCAAAAAATCAAAAAAT--------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 772
E14             AATCAAAAAATCAAAAAAT--------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 772
E15             AATCAAAAAATCAAAAAAT--------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 772
E17             AATCAAAAAAT----------------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 764
E20             AATCAAAAAATCAAAAAATCAAAAAATACGAAGATTTTACAACCATTGTGCAGAAGAACC 780
E22             AATCAAAAAAT----------------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 764
E23             AATCAAAAAATCAAAAAAT--------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 772
E24             AATCAAAAAATCAAAAAAT--------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 772
E25             AATCAAAAAATCAAAAAATCAAAAAATACGAAGATTTTACAACCATTGTGCAGAAGAACC 780
E26             AATCAAAAAATCAAAAAAT--------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 772
E27             AATCAAAAAATCAAAAAAT--------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 772
E28             AATCAAAAAAT----------------ACGAAGATTTTACAACCATTGTGCAGAAGAACC 764
                ***********                **********************************

WT3358          CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 832
E1              CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 832
E2              CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 832
E3              CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 824
E4              CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 832
E7              CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 832
E8              CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 832
E10             CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 840
E11             CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 832
E12             CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 832
E14             CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 832
E15             CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 832
E17             CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 824
E20             CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 840
E22             CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 824
E23             CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 832
E24             CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 832
E25             CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 840
E26             CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 832
E27             CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 832
E28             CCGAAAAGCCGTTAACGTTCATAGTTGAGCGTAACGGCAAAGAAGAGCAACTAACAGTGA 824
                ************************************************************

WT3358          CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 892
E1              CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 892
E2              CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 892
E3              CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 884
E4              CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 892
E7              CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 892
E8              CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 892
E10             CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 900
E11             CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 892
E12             CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 892
E14             CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 892
E15             CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 892
E17             CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 884
E20             CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 900
E22             CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 884
E23             CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 892
E24             CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 892
E25             CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 900
E26             CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 892
E27             CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 892
E28             CACCAGAAAAACAAAAAGTGGAAAAACAAACGATTGGTAAAGTCGGCGTTTATCCTTATA 884
                ************************************************************

WT3358          TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 952
E1              TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 952
E2              TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 952
E3              TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 944
E4              TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 952
E7              TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 952
E8              TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 952
E10             TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 960
E11             TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 952
E12             TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 952
E14             TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 952
E15             TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 952
E17             TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 944
E20             TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 960
E22             TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 944
E23             TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 952
```

```
E24             TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 952
E25             TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 960
E26             TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 952
E27             TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 952
E28             TGAAAACCGATTTACCGTCAAAATTGATGGGCGGTATTCAGGATACTTTGAATAGTACGA 944
                ************************************************************

WT3358          CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1012
E1              CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1012
E2              CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1012
E3              CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1004
E4              CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1012
E7              CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1012
E8              CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1012
E10             CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1020
E11             CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1012
E12             CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1012
E14             CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1012
E15             CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1012
E17             CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1004
E20             CACAGATTTTTAAAGC-CTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1019
E22             CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1004
E23             CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1012
E24             CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1012
E25             CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1020
E26             CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1012
E27             CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1012
E28             CACAGATTTTTAAAGCACTCGGCTCACTATTCACAGGCTTTAGTTTAAACAAACTAGGTG 1004
                *************** ******************************************

WT3358          GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1072
E1              GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1072
E2              GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1072
E3              GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1064
E4              GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1072
E7              GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1072
E8              GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1072
E10             GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1080
E11             GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1072
E12             GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1072
E14             GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1072
E15             GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1072
E17             GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1064
E20             GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1079
E22             GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1064
E23             GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1072
E24             GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1072
E25             GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1080
E26             GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1072
E27             GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1072
E28             GGCCAGTCATGATGTTTAAATTATCGGAAGAAGCATCCAATGCTGGAGTAAGTACAGTTG 1064
                ************************************************************

WT3358          TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1132
E1              TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1132
E2              TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1132
E3              TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1124
E4              TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1132
E7              TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1132
E8              TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1132
E10             TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1140
E11             TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1132
E12             TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1132
E14             TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1132
E15             TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1132
E17             TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1124
E20             TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1139
E22             TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1124
E23             TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1132
E24             TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1132
E25             TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1140
E26             TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1132
E27             TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1132
E28             TATTCTTAATGGCCATGTTGTCAATGAACTTAGGGATTATTAATTTGTTGCCAATCCCAG 1124
                ************************************************************

WT3358          CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1192
E1              CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1192
E2              CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1192
E3              CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1184
E4              CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1192
E7              CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1192
E8              CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1192
E10             CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1200
E11             CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1192
E12             CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1192
E14             CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1192
E15             CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1192
E17             CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1184
E20             CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1199
E22             CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1184
E23             CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1192
```

```
E24             CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1192
E25             CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1200
E26             CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1192
E27             CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1192
E28             CTTTAGATGGCGGGAAAATTGTCTTAAACATTATTGAAGGTGTACGTGGAAAACCAATTA 1184
                ************************************************************

WT3358          GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1252
E1              GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1252
E2              GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1252
E3              GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1244
E4              GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1252
E7              GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1252
E8              GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1252
E10             GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1260
E11             GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1252
E12             GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1252
E14             GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1252
E15             GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1252
E17             GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1244
E20             GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1259
E22             GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1244
E23             GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1252
E24             GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1252
E25             GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1260
E26             GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1252
E27             GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1252
E28             GTCCTGAAAAAGAAGGCATCATTACGTTAATTGGCTTTGGGTTTGTCATGGTGTTAATGG 1244
                ************************************************************

WT3358          TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1293
E1              TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1293
E2              TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1293
E3              TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1285
E4              TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1293
E7              TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1293
E8              TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1293
E10             TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1301
E11             TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1293
E12             TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1293
E14             TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1293
E15             TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1293
E17             TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1285
E20             TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1300
E22             TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1285
E23             TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1293
E24             TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1294
E25             TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1301
E26             TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1293
E27             TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1293
E28             TGTTAGTTACTTGGAACGATATTCAACGCTTTTTCTTTTAA 1285
                ****************************************
```

A2: MSA of 3358 RseP

```
WT3358          MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E1              MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E2              MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E3              MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E4              MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E7              MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E8              MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E10             MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E11             MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E12             MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E14             MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E15             MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E17             MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E20             MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E22             MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E23             MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E24             MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E25             MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E26             MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E27             MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
E28             MKEERRTFMKTIITFIIVFGILVLVHEFGHFYFAKRAGILVREFAIGMGPKIFAHRGKDG 60
                ************************************************************

WT3358          TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E1              TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E2              TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E3              TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E4              TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E7              TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E8              TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E10             TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E11             TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E12             TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E14             TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
```

```
E15             TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E17             TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E20             TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E22             TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E23             TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E24             TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E25             TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E26             TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E27             TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
E28             TTYTIRLLPIGGYVRMAGMGEDMTEITPGMPLSVELNAVGNVVKINTSKKVQLPHSIPME 120
                ************************************************************

WT3358          VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E1              VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E2              VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E3              VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E4              VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E7              VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E8              VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E10             VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E11             VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E12             VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E14             VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E15             VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E17             VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E20             VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E22             VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E23             VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E24             VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E25             VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E26             VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E27             VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
E28             VVDFDLEKELFIKGYVNGNEEEETVYKVDHDATIIESDGTEVRIAPLDVQFQSAKLSQRI 180
                ************************************************************

WT3358          LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E1              LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E2              LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E3              LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E4              LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E7              LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E8              LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E10             LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E11             LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E12             LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E14             LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E15             LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E17             LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E20             LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E22             LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E23             LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E24             LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E25             LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E26             LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E27             LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
E28             LTNFAGPMNNFILGFILFTLAVFLQGGVTDLNTNQIGQVIPNGPAAEAGLKENDKVLSIN 240
                ************************************************************

WT3358          NQKIKKYEDFTTIVQKNPEKPLTFIVERNGKEEQLTVTPEKQKVEKQTIGKVGVYPYMKT 300
E1              NQKIKKYEDFTTIVQKNPEKPLTFIVERNGKEEQLTVTPEKQKVEKQTIGKVGVYPYMKT 300
E2              NQKIKKYEDFTTIVQKNPEKPLTFIVERNGKEEQLTVTPEKQKVEKQTIGKVGVYPYMKT 300
E3              NQKIRRFYNHCAEEPRKAVNVHS------------------------------------- 263
E4              NQKIKKYEDFTTIVQKNPEKPLTFIVERNGKEEQLTVTPEKQKVEKQTIGKVGVYPYMKT 300
E7              NQKIKKYEDFTTIVQKNPEKPLTFIVERNGKEEQLTVTPEKQKVEKQTIGKVGVYPYMKT 300
E8              NQKIKKYEDFTTIVQKNPEKPLTFIVERNGKEEQLTVTPEKQKVEKQTIGKVGVYPYMKT 300
E10             NQKIKKSKNTKILQPLCRRTPKSR---------------------------------- 264
E11             NQKIKKYEDFTTIVQKNPEKPLTFIVERNGKEEQLTVTPEKQKVEKQTIGKVGVYPYMKT 300
E12             NQKIKKYEDFTTIVQKNPEKPLTFIVERNGKEEQLTVTPEKQKVEKQTIGKVGVYPYMKT 300
E14             NQKIKKYEDFTTIVQKNPEKPLTFIVERNGKEEQLTVTPEKQKVEKQTIGKVGVYPYMKT 300
E15             NQKIKKYEDFTTIVQKNPEKPLTFIVERNGKEEQLTVTPEKQKVEKQTIGKVGVYPYMKT 300
E17             NQKIRRFYNHCAEEPRKAVNVHS------------------------------------- 263
E20             NQKIKKSKNTKILQPLCRRTPKSR---------------------------------- 264
E22             NQKIRRFYNHCAEEPRKAVNVHS------------------------------------- 263
E23             NQKIKKYEDFTTIVQKNPEKPLTFIVERNGKEEQLTVTPEKQKVEKQTIGKVGVYPYMKT 300
E24             NQKIKKYEDFTTIVQKNPEKPLTFIVERNGKEEQLTVTPEKQKVEKQTIGKVGVYPYMKT 300
E25             NQKIKKSKNTKILQPLCRRTPKSR---------------------------------- 264
E26             NQKIKKYEDFTTIVQKNPEKPLTFIVERNGKEEQLTVTPEKQKVEKQTIGKVGVYPYMKT 300
E27             NQKIKKYEDFTTIVQKNPEKPLTFIVERNGKEEQLTVTPEKQKVEKQTIGKVGVYPYMKT 300
E28             NQKIRRFYNHCAEEPRKAVNVHS------------------------------------- 263
                ****::  :                     : : :

WT3358          DLPSKLMGGIQDTLNSTTQIFKALGSLFTGFSLNKLGGPVMMFKLSEEASNAGVSTVVFL 360
E1              DLPSKLMGGIQDTLNSTTQIFKALGSLFTGFSLNKLGGPVMMFKLSEEASNAGVSTVVFL 360
E2              DLPSKLMGGIQDTLNSTTQIFKALGSLFTGFSLNKLGGPVMMFKLSEEASNAGVSTVVFL 360
E3              ------------------------------------------------------------
E4              DLPSKLMGGIQDTLNSTTQIFKALGSLFTGFSLNKLGGPVMMFKLSEEASNAGVSTVVFL 360
E7              DLPSKLMGGIQDTLNSTTQIFKALGSLFTGFSLNKLGGPVMMFKLSEEASNAGVSTVVFL 360
E8              DLPSKLMGGIQDTLNSTTQIFKALGSLFTGFSLNKLGGPVMMFKLSEEASNAGVSTVVFL 360
E10             ------------------------------------------------------------
E11             DLPSKLMGGIQDTLNSTTQIFKALGSLFTGFSLNKLGGPVMMFKLSEEASNAGVSTVVFL 360
E12             DLPSKLMGGIQDTLNSTTQIFKALGSLFTGFSLNKLGGPVMMFKLSEEASNAGVSTVVFL 360
E14             DLPSKLMGGIQDTLNSTTQIFKALGSLFTGFSLNKLGGPVMMFKLSEEASNAGVSTVVFL 360
```

```
E15             DLPSKLMGGIQDTLNSTTQIFKALGSLFTGFSLNKLGGPVMMFKLSEEASNAGVSTVVFL 360
E17             ------------------------------------------------------------
E20             ------------------------------------------------------------
E22             ------------------------------------------------------------
E23             DLPSKLMGGIQDTLNSTTQIFKALGSLFTGFSLNKLGGPVMMFKLSEEASNAGVSTVVFL 360
E24             DLPSKLMGGIQDTLNSTTQIFKALGSLFTGFSLNKLGGPVMMFKLSEEASNAGVSTVVFL 360
E25             ------------------------------------------------------------
E26             DLPSKLMGGIQDTLNSTTQIFKALGSLFTGFSLNKLGGPVMMFKLSEEASNAGVSTVVFL 360
E27             DLPSKLMGGIQDTLNSTTQIFKALGSLFTGFSLNKLGGPVMMFKLSEEASNAGVSTVVFL 360
E28             ------------------------------------------------------------

WT3358          MAMLSMNLGIINLLPIPALDGGKIVLNIIEGVRGKPISPEKEGIITLIGFGFVMVLMVLVTWNDIQRFFF 430
E1              MAMLSMNLGIINLLPIPALDGGKIVLNIIEGVRGKPISPEKEGIITLIGFGFVMVLMVLVTWNDIQRFFF 430
E2              MAMLSMNLGIINLLPIPALDGGKIVLNIIEGVRGKPISPEKEGIITLIGFGFVMVLMVLVTWNDIQRFFF 430
E3              ------------------------------------------------------------------------
E4              MAMLSMNLGIINLLPIPALDGGKIVLNIIEGVRGKPISPEKEGIITLIGFGFVMVLMVLVTWNDIQRFFF 430
E7              MAMLSMNLGIINLLPIPALDGGKIVLNIIEGVRGKPISPEKEGIITLIGFGFVMVLMVLVTWNDIQRFFF 430
E8              MAMLSMNLGIINLLPIPALDGGKIVLNIIEGVRGKPISPEKEGIITLIGFGFVMVLMVLVTWNDIQRFFF 430
E10             ------------------------------------------------------------------------
E11             MAMLSMNLGIINLLPIPALDGGKIVLNIIEGVRGKPISPEKEGIITLIGFGFVMVLMVLVTWNDIQRFFF 430
E12             MAMLSMNLGIINLLPIPALDGGKIVLNIIEGVRGKPISPEKEGIITLIGFGFVMVLMVLVTWNDIQRFFF 430
E14             MAMLSMNLGIINLLPIPALDGGKIVLNIIEGVRGKPISPEKEGIITLIGFGFVMVLMVLVTWNDIQRFFF 430
E15             MAMLSMNLGIINLLPIPALDGGKIVLNIIEGVRGKPISPEKEGIITLIGFGFVMVLMVLVTWNDIQRFFF 430
E17             ------------------------------------------------------------------------
E20             ------------------------------------------------------------------------
E22             ------------------------------------------------------------------------
E23             MAMLSMNLGIINLLPIPALDGGKIVLNIIEGVRGKPISPEKEGIITLIGFGFVMVLMVLVTWNDIQRFFF 430
E24             MAMLSMNLGIINLLPIPALDGGKIVLNIIEGVRGKPISPEKEGIITLIGFGFVMVLMVLVTWNDIQRFFF 430
E25             ------------------------------------------------------------------------
E26             MAMLSMNLGIINLLPIPALDGGKIVLNIIEGVRGKPISPEKEGIITLIGFGFVMVLMVLVTWNDIQRFFF 430
E27             MAMLSMNLGIINLLPIPALDGGKIVLNIIEGVRGKPISPEKEGIITLIGFGFVMVLMVLVTWNDIQRFFF 430
E28             ------------------------------------------------------------------------
```

Pediococcus acidilactici 2002

A3: MSA of 2002 *rseP*

```
2002WT          ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M1              ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M2              ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M3              ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M4              ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M5              ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M6              ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M7              ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M8              ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M9              ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M10             ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M11             ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M12             ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M13             ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M14             ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M15             ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M16             ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M17             ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M18             ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M19             ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M20             ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M21             ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
M22             ATGATTACTACGGTAATTACTTTTTTGATTGTTTTTTGCATTTTAGTTATCGTCCATGAA 60
                ************************************************************

2002WT          TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M1              TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M2              TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M3              TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M4              TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M5              TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M6              TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M7              TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M8              TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M9              TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M10             TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M11             TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M12             TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M13             TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M14             TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M15             TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M16             TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M17             TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M18             TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M19             TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M20             TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M21             TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
M22             TACGGACATTTTTTGGCTGCAAAAAAGTCCGGCATTTTAGTCCGTGAATTTTCAGTAGGT 120
                ************************************************************
```

```
2002WT          ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M1              ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M2              ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M3              ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M4              ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M5              ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M6              ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M7              ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M8              ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M9              ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M10             ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M11             ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M12             ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M13             ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M14             ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M15             ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M16             ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M17             ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M18             ATGTGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M19             ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M20             ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M21             ATGTGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
M22             ATGGGACCGAAGATTGTTGATCTTAAACGACGCGGAACCACGTTTACCCTGCGAATTCTC 180
                *** ************************************************************

2002WT          CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M1              CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M2              CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M3              CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M4              CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M5              CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M6              CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M7              CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M8              CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M9              CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M10             CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M11             CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M12             CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M13             CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M14             CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M15             CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M16             CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M17             CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M18             CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M19             CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M20             CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M21             CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
M22             CCAATCGGCGGCTACGTACGAATGGCGGGCTTGGATGAAGAAGACGAGGAATTAAAAGCG 240
                ************************************************************

2002WT          GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M1              GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M2              GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M3              GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M4              GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M5              GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M6              GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M7              GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M8              GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M9              GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M10             GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M11             GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M12             GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M13             GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M14             GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M15             GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M16             GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M17             GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M18             GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M19             GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M20             GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M21             GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
M22             GGACAGCCGGTAGTGTTAACAACCAATTCTGCGGGAGTAGTAACTACAATCAATACTTCT 300
                ************************************************************

2002WT          GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M1              GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M2              GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M3              GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M4              GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M5              GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M6              GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M7              GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M8              GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M9              GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M10             GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M11             GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M12             GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M13             GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M14             GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M15             GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
```

```
M16             GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M17             GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M18             GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M19             GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M20             GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M21             GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
M22             GAGAAGTTTCGTTCTGCCGTCGGTGTACCGTTAATCGTTACTAAATTTGATCTGCAAGAC 360
                ************************************************************

2002WT          GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M1              GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M2              GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M3              GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M4              GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M5              GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M6              GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M7              GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M8              GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M9              GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M10             GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M11             GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M12             GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M13             GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M14             GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M15             GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M16             GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M17             GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M18             GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M19             GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M20             GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M21             GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
M22             GCACTCTTCATCGAAGGTTACGAAAACGGCGATGAAGGTGCAACCAAACGTTTTGCAGTC 420
                ************************************************************

2002WT          GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M1              GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M2              GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M3              GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M4              GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M5              GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M6              GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M7              GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M8              GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M9              GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M10             GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M11             GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M12             GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M13             GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M14             GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M15             GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M16             GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M17             GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M18             GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M19             GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M20             GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M21             GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
M22             GATCACGATGCAACGATTGTTGAAAAAGATGGAACCGAAGTACGAATCGCTCCCCGTGAC 480
                ************************************************************

2002WT          GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M1              GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M2              GTACAGTTCCAATCAGCATCGGTTTGGCGCAAACTACTAACTAATTTTGCTGGGCCATTT 540
M3              GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M4              GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M5              GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M6              GTACA-TTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 539
M7              GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M8              GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M9              GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M10             GTACAGTTCCAATCAGCATCGGTTTGGCGCAAACTACTAACTAATTTTGCTGGGCCATTT 540
M11             GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M12             GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M13             GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M14             GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M15             GTACAGTTCCAATCAGCATCGGTTTGGCGCAAACTACTAACTAATTTTGCTGGGCCATTT 540
M16             GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M17             GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M18             GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M19             GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M20             GTACAGTTCCAATCAGCATCGGTTTAGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M21             GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
M22             GTACAGTTCCAATCAGCATCGGTTTGGCGCAGACTACTAACTAATTTTGCTGGGCCATTT 540
                ***** ******************.*****.****************************

2002WT          AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M1              AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M2              AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M3              AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M4              AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M5              AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M6              AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 599
M7              AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
```

```
M8              AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M9              AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M10             AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M11             AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M12             AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M13             AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M14             AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M15             AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M16             AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M17             AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M18             AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M19             AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M20             AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M21             AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
M22             AATAATTTCATCTTAGCCATCGTGGTTTTTGCGTTAATGGGGATCTTGCAGGGCGCGGTG 600
                ************************************************************

2002WT          CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M1              CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M2              CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M3              CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M4              CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M5              CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M6              CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 659
M7              CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M8              CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M9              CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M10             CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M11             CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M12             CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M13             CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M14             CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M15             CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M16             CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M17             CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M18             CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M19             CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M20             CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M21             CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
M22             CCTTCAAATTCAAATCAAGTTCAGGTAATTGATAACGGAGTGGCTCAAAAAGCGGGCGTT 660
                ************************************************************

2002WT          AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M1              AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M2              AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M3              AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M4              AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M5              AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M6              AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 719
M7              AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M8              AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M9              AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M10             AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M11             AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M12             AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M13             AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M14             AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M15             AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M16             AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M17             AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M18             AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M19             AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M20             AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M21             AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
M22             AGGAATAACGACCGGATCGTTGCTGTAGATGGTCAAAAGACGCAGAATTGGTCGGCAATT 720
                ************************************************************

2002WT          TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M1              TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M2              TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M3              TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M4              TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M5              TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M6              TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 779
M7              TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M8              TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M9              TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M10             TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M11             TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M12             TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M13             TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M14             TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M15             TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M16             TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M17             TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M18             TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M19             TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M20             TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M21             TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
M22             TCAAAAGCCGTTTCAAGCCACCCTAAACAGTCGATAACGCTTAAGCTGCAAAAAAATGGG 780
                ************************************************************
```

```
2002WT          AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M1              AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAGG-TCGGA 839
M2              AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M3              AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M4              AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M5              AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M6              AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 839
M7              AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M8              AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M9              AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M10             AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M11             AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M12             AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M13             AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M14             AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M15             AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M16             AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M17             AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M18             AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M19             AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M20             AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M21             AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
M22             AAGACTCGTTCGGTGCGGGTTACCCCTAAGGCGGTTAATAATGGTCAAAAAAAGGTCGGA 840
                ************************************************.* *****

2002WT          ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M1              ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 899
M2              ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M3              ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M4              ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M5              ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M6              ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 899
M7              ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M8              ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M9              ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M10             ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M11             ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M12             ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M13             ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M14             ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M15             ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M16             ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M17             ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M18             ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M19             ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M20             ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M21             ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
M22             ATGATTGGGATTCAATCATCAATGACCACTAATCTGGGTTCTCGAATTATGTACGGTTTT 900
                ************************************************************

2002WT          ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M1              ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 959
M2              ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M3              ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M4              ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M5              ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M6              ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 959
M7              ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M8              ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M9              ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M10             ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M11             ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M12             ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M13             ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M14             ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M15             ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M16             ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M17             ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M18             ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M19             ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M20             ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M21             ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
M22             ACGGGGACTTGGCAAATGACTAAGGCTTTATTTAGTGCTTTAGCACAAATGCTGCACGGC 960
                ************************************************************

2002WT          TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M1              TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1019
M2              TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M3              TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M4              TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M5              TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M6              TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1019
M7              TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M8              TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M9              TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M10             TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M11             TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M12             TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M13             TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M14             TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M15             TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M16             TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
```

```
M17             TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M18             TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M19             TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M20             TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M21             TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
M22             TTTAGTTTAAATGATCTGGGTGGACCAGTAGCAATCTACGCAACCACGTCAAAAGCTACC 1020
                ************************************************************

2002WT          CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M1              CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1079
M2              CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M3              CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M4              CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M5              TAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M6              CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1079
M7              CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M8              CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M9              CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M10             CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAAACTTGGAATC 1080
M11             CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M12             CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M13             CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M14             CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M15             CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M16             CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M17             CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M18             CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M19             CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M20             CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M21             CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
M22             CAACAAGGCTTCAATGCAGTTTTATATATTCTGGGATTCCTATCCTTAAACCTTGGAATC 1080
                 ************************************************.*********

2002WT          GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M1              GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1139
M2              GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M3              GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M4              GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M5              GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M6              GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1139
M7              GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M8              GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M9              GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M10             GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M11             GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M12             GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M13             GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M14             GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M15             GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M16             GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M17             GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M18             GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M19             GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M20             GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M21             GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
M22             GTCAATTTATTGCCCATTCCAGCGTTAGATGGCGGAAAGATTTTATTAAACTTAATTGAA 1140
                ************************************************************

2002WT          GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M1              GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1199
M2              GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M3              GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M4              GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M5              GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M6              GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1199
M7              GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M8              GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M9              GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M10             GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M11             GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M12             GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M13             GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M14             GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M15             GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M16             GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M17             GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M18             GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M19             GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M20             GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M21             GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
M22             GTGGTTCGGCGTAAGCCAATGAAAATGGAGACAGAGAATATGATTACACTAGTCGGCTTT 1200
                ************************************************************

2002WT          GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M1              GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1259
M2              GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M3              GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M4              GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M5              GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M6              GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1259
M7              GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M8              GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
```

```
M9              GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M10             GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M11             GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M12             GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M13             GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M14             GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M15             GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M16             GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M17             GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M18             GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M19             GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M20             GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M21             GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
M22             GGTTTCTTAATGTTGCTAATGCTGCTAGTTACTTGGAACGATATTCAAAGATACTTTTTC 1260
                ************************************************************
```

A4: MSA of 2002 RseP

```
2002WT          MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M1              MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M2              MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M3              MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M4              MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M5              MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M6              MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M7              MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M8              MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M9              MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M10             MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M11             MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M12             MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M13             MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M14             MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M15             MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M16             MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M17             MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M18             MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGM------------------- 41
M19             MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M20             MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
M21             MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGM------------------- 41
M22             MITTVITFLIVFCILVIVHEYGHFLAAKKSGILVREFSVGMGPKIVDLKRRGTTFTLRIL 60
                ************************************

2002WT          PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M1              PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M2              PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M3              PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M4              PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M5              PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M6              PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M7              PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M8              PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M9              PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M10             PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M11             PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M12             PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M13             PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M14             PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M15             PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M16             PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M17             PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M18             ------------------------------------------------------------ 
M19             PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M20             PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120
M21             ------------------------------------------------------------ 
M22             PIGGYVRMAGLDEEDEELKAGQPVVLTTNSAGVVTTINTSEKFRSAVGVPLIVTKFDLQD 120

2002WT          ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
M1              ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
M2              ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRKLLTNFAGPF 180
M3              ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
M4              ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
M5              ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
M6              ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVHSNQHRFGADY-------- 172
M7              ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
M8              ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
M9              ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
M10             ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRKLLTNFAGPF 180
M11             ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
M12             ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
M13             ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
M14             ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
M15             ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRKLLTNFAGPF 180
M16             ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
M17             ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
M18             ------------------------------------------------------------ 
M19             ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180
```

```
M20             ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASV------------ 168
M21             ------------------------------------------------------------
M22             ALFIEGYENGDEGATKRFAVDHDATIVEKDGTEVRIAPRDVQFQSASVWRRLLTNFAGPF 180

2002WT          NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M1              NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M2              NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M3              NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M4              NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M5              NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M6              ------------------------------------------------------------
M7              NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M8              NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M9              NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M10             NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M11             NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M12             NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M13             NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M14             NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M15             NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M16             NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M17             NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M18             ------------------------------------------------------------
M19             NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240
M20             ------------------------------------------------------------
M21             ------------------------------------------------------------
M22             NNFILAIVVFALMGILQGAVPSNSNQVQVIDNGVAQKAGVRNNDRIVAVDGQKTQNWSAI 240

2002WT          SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M1              SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKRSE-------------------- 280
M2              SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M3              SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M4              SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M5              SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M6              ------------------------------------------------------------
M7              SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M8              SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M9              SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M10             SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M11             SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M12             SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M13             SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M14             SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M15             SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M16             SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M17             SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M18             ------------------------------------------------------------
M19             SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300
M20             ------------------------------------------------------------
M21             ------------------------------------------------------------
M22             SKAVSSHPKQSITLKLQKNGKTRSVRVTPKAVNNGQKKVGMIGIQSSMTTNLGSRIMYGF 300

2002WT          TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M1              ------------------------------------------------------------
M2              TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M3              TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M4              TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M5              TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKAT---------------------- 340
M6              ------------------------------------------------------------
M7              TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M8              TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M9              TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M10             TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLKLGI 360
M11             TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M12             TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M13             TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M14             TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M15             TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M16             TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M17             TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M18             ------------------------------------------------------------
M19             TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360
M20             ------------------------------------------------------------
M21             ------------------------------------------------------------
M22             TGTWQMTKALFSALAQMLHGFSLNDLGGPVAIYATTSKATQQGFNAVLYILGFLSLNLGI 360

2002WT          VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M1              ------------------------------------------------------------
M2              VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M3              VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M4              VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M5              ------------------------------------------------------------
M6              ------------------------------------------------------------
M7              VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M8              VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M9              VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
```

```
M10             VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M11             VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M12             VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M13             VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M14             VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M15             VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M16             VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M17             VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M18             ------------------------------------------------------------
M19             VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
M20             ------------------------------------------------------------
M21             ------------------------------------------------------------
M22             VNLLPIPALDGGKILLNLIEVVRRKPMKMETENMITLVGFGFLMLLMLLVTWNDIQRYFF 420
```

Lactococcus garvieae DCC43

A5. MSA of DCC43 *rseP*

```
wtDCC43         ATGGGGCCGAAAATTTTTGCGCATCAAGCCAAAGACGGTACATTTTATAGTATTCGTATT 60
G1              ATGGGGCCGAAAATTTTTGCGCATCAAGCCAAAGACGGTACATTTTATAGTATTCGTATT 60
G2              ATGGGGCCGAAAATTTTTGCGCATCAAGCCAAAGACGGTACATTTTATAGTATTCGTATT 60
G3              ATGGGGCCGAAAATTTTTGCGCATCAAGCCAAAGACGGTACATTTTATAGTATTCGTATT 60
G4              ATGGGGCCGAAAATTTTTGCGCATCAAGCCAAAGACGGTACATTTTATAGTATTCGTATT 60
G5              ATGGGGCCGAAAATTTTTGCGCATCAAGCCAAAGACGGTACATTTTATAGTATTCGTATT 60
G6              ATGGGGCCGAAAATTTTTGCGCATCAAGCCAAAGACGGTACATTTTATAGTATTCGTATT 60
G7              ATGGGGCCGAAAATTTTTGCGCATCAAGCCAAAGACGGTACATTTTATAGTATTCGTATT 60
G8              ATGGGGCCGAAAATTTTTGCGCATCAAGCCAAAGACGGTACATTTTATAGTATTCGTATT 60
G9              ATGGGGCCGAAAATTTTTGCGCATCAAGCCAAAGACGGTACATTTTATAGTATTCGTATT 60
G10             ATGGGGCCGAAAATTTTTGCGCATCAAGCCAAAGACGGTACATTTTATAGTATTCGTATT 60
                ************************************************************

wtDCC43         CTTCCTCTTGGTGGTTATGTTCGTATGGCTGGCTGGGGTGACGATACTACAGAAATAAAA 120
G1              CTTCCTCTTGGTGGTTATGTTCGTATGGCTGGCTGGGGTGACGATACTACAGAAATAAAA 120
G2              CTTCCTCTTGGTGGTTATGTTCGTATGGCTGGCTGGGGTGACGATACTACAGAAATAAAA 120
G3              CTTCCTCTTGGTGGTTATGTTCGTATGGCTGGCTGGGGTGACGATACTACAGAAATAAAA 120
G4              CTTCCTCTTGGTGGTTATGTTCGTATGGCTGGCTGGGGTGACGATACTACAGAAATAAAA 120
G5              CTTCCTCTTGGTGGTTATGTTCGTATGGCTGGCTGGGGTGACGATACTACAGAAATAAAA 120
G6              CTTCCTCTTGGTGGTTATGTTCGTATGGCTGGCTGGGGTGACGATACTACAGAAATAAAA 120
G7              CTTCCTCTTGGTGGTTATGTTCGTATGGCTGGCTGGGGTGACGATACTACAGAAATAAAA 120
G8              CTTCCTCTTGGTGGTTATGTTCGTATGGCTGGCTGGGGTGACGATACTACAGAAATAAAA 120
G9              CTTCCTCTTGGTGGTTATGTTCGTATGGCTGGCTGGGGTGACGATACTACAGAAATAAAA 120
G10             CTTCCTCTTGGTGGTTATGTTCGTATGGCTGGCTGGGGTGACGATACTACAGAAATAAAA 120
                ************************************************************

wtDCC43         AAAGGTACACCAGCCTCTGTAATGATTGTTGACAATCAGGTCGTAAAAATTAATCTTTCA 180
G1              AAAGGTACACCAGCCTCTGTAATGATTGTTGACAATCAGGTCGTAAAAATTAATCTTTCA 180
G2              AAAGGTACACCAGCCTCTGTAATGATTGTTGACAATCAGGTCGTAAAAATTAATCTTTCA 180
G3              AAAGGTACACCAGCCTCTGTAATGATTGTTGACAATCAGGTCGTAAAAATTAATCTTTCA 180
G4              AAAGGTACACCAGCCTCTGTAATGATTGTTGACAATCAGGTCGTAAAAATTAATCTTTCA 180
G5              AAAGGTACACCAGCCTCTGTAATGATTGTTGACAATCAGGTCGTAAAAATTAATCTTTCA 180
G6              AAAGGTACACCAGCCTCTGTAATGATTGTTGACAATCAGGTCGTAAAAATTAATCTTTCA 180
G7              AAAGGTACACCAGCCTCTGTAATGATTGTTGACAATCAGGTCGTAAAAATTAATCTTTCA 180
G8              AAAGGTACACCAGCCTCTGTAATGATTGTTGACAATCAGGTCGTAAAAATTAATCTTTCA 180
G9              AAAGGTACACCAGCCTCTGTAATGATTGTTGACAATCAGGTCGTAAAAATTAATCTTTCA 180
G10             AAAGGTACACCAGCCTCTGTAATGATTGTTGACAATCAGGTCGTAAAAATTAATCTTTCA 180
                ************************************************************

wtDCC43         GACCGGGTGACCTTGGAAAATAGTGTCCCCATGTTGGTCACTGAATACGATTTTGAAAAA 240
G1              GACCGGGTGACCTTGGAAAATAGTGTCCCCATGTTGGTCACTGAATACGATTTTGAAAAA 240
G2              GACCGGGTGACCTTGGAAAATAGTGTCCCCATGTTGGTCACTGAATACGATTTTGAAAAA 240
G3              GACCGGGTGACCTTGGAAAATAGTGTCCCCATGTTGGTCACTGAATACGATTTTGAAAAA 240
G4              GACCGGGTGACCTTGGAAAATAGTGTCCCCATGTTGGTCACTGAATACGATTTTGAAAAA 240
G5              GACCGGGTGACCTTGGAAAATAGTGTCCCCATGTTGGTCACTGAATACGATTTTGAAAAA 240
G6              GACCGGGTGACCTTGGAAAATAGTGTCCCCATGTTGGTCACTGAATACGATTTTGAAAAA 240
G7              GACCGGGTGACCTTGGAAAATAGTGTCCCCATGTTGGTCACTGAATACGATTTTGAAAAA 240
G8              GACCGGGTGACCTTGGAAAATAGTGTCCCCATGTTGGTCACTGAATACGATTTTGAAAAA 240
G9              GACCGGGTGACCTTGGAAAATAGTGTCCCCATGTTGGTCACTGAATACGATTTTGAAAAA 240
G10             GACCGGGTGACCTTGGAAAATAGTGTCCCCATGTTGGTCACTGAATACGATTTTGAAAAA 240
                ************************************************************

wtDCC43         GAACTTTTCATCGAAGGCGAAGTTTTTGGAGAAATTACACGTTATTCAGTTGACCATGAT 300
G1              GAACTTTTCATCGAAGGCGAAGTTTTTGGAGAAATTACACGTTATTCAGTTGACCATGAT 300
G2              GAACTTTTCATCGAAGGCGAAGTTTTTGGAGAAATTACACGTTATTCAGTTGACCATGAT 300
G3              GAACTTTTCATCGAAGGCGAAGTTTTTGGAGAAATTACACGTTATTCAGTTGACCATGAT 300
G4              GAACTTTTCATCGAAGGCGAAGTTTTTGGAGAAATTACACGTTATTCAGTTGACCATGAT 300
G5              GAACTTTTCATCGAAGGCGAAGTTTTTGGAGAAATTACACGTTATTCAGTTGACCATGAT 300
G6              GAACTTTTCATCGAAGGCGAAGTTTTTGGAGAAATTACACGTTATTCAGTTGACCATGAT 300
G7              GAACTTTTCATCGAAGGCGAAGTTTTTGGAGAAATTACACGTTATTCAGTTGACCATGAT 300
G8              GAACTTTTCATCGAAGGCGAAGTTTTTGGAGAAATTACACGTTATTCAGTTGACCATGAT 300
G9              GAACTTTTCATCGAAGGCGAAGTTTTTGGAGAAATTACACGTTATTCAGTTGACCATGAT 300
G10             GAACTTTTCATCGAAGGCGAAGTTTTTGGAGAAATTACACGTTATTCAGTTGACCATGAT 300
                ************************************************************

wtDCC43         GCAACGATCATTGAAGGCGATGGCACGGAAATCCGTATCGCGCCTCTCGACGTACAATAT 360
G1              GCAACGATCATTGAAGGCGATGGCACGGAAATCCGTATCGCGCCTCTCGACGTACAATAT 360
G2              GCAACGATCATTGAAGGCGATGGCACGGAAATCCGTATCGCGCCTCTCGACGTACAATAT 360
G3              GCAACGATCATTGAAGGCGATGGCACGGAAATCCGTATCGCGCCTCTCGACGTACAATAT 360
```

```
G4              GCAACGATCATTGAAGGCGATGGCACGGAAATCCGTATCGCGCCTCTCGACGTACAATAT 360
G5              GCAACGATCATTGAAGGCGATGGCACGGAAATCCGTATCGCGCCTCTCGACGTACAATAT 360
G6              GCAACGATCATTGAAGGCGATGGCACGGAAATCCGTATCGCGCCTCTCGACGTACAATAT 360
G7              GCAACGATCATTGAAGGCGATGGCACGGAAATCCGTATCGCGCCTCTCGACGTATAATAT 360
G8              GCAACGATCATTGAAGGCGATGGCACGGAAATCCGTATCGCGCCTCTCGACGTACAATAT 360
G9              GCAACGATCATTGAAGGCGATGGCACGGAAATCCGTATCGCGCCTCTCGACGTACAATAT 360
G10             GCAACGATCATTGAAGGCGATGGCACGGAAATCCGTATCGCGCCTCTCGACGTACAATAT 360
                ************************************************** *****

wtDCC43         CAAAGTGCTAGCGTAATCGGCAAGATATTGACAAACTTTGCAGGTCCTTTAAATAATTTT 420
G1              CAAAGTGCTAGCGTAATCGGCAAGATATTGACAAACTTTGCAGGTCCTTTAAATAATTTT 420
G2              CAAAGTGCTAGCGTAATCGGCAAGATATTGACAAACTTTGCAGGTCCTTTAAATAATTTT 420
G3              CAAAGTGCTAGCGTAATCGGCAAGATATTGACAAACTTTGCAGGTCCTTTAAATAATTTT 420
G4              CAAAGTGCTAGCGTAATCGGCAAGATATTGACAAACTTTGCAGGTCCTTTAAATAATTTT 420
G5              CAAAGTGCTAGCGTAATCGGCAAGATATTGACAAACTTTGCAGGTCCTTTAAATAATTTT 420
G6              CAAAGTGCTAGCGTAATCGGCAAGATATTGACAAACTTTGCAGGTCCTTTAAATAATTTT 420
G7              CAAAGTGCTAGCGTAATCGGCAAGATATTGACAAACTTTGCAGGTCCTTTAAATAATTTT 420
G8              CAAAGTGCTAGCGTAATCGGCAAGATATTGACAAACTTTGCAGGTCCTTTAAATAATTTT 420
G9              CAAAGTGCTAGCGTAATCGGCAAGATATTGACAAACTTTGCAGGTCCTTTAAATAATTTT 420
G10             CAAAGTGCTAGCGTAATCGGCAAGATATTGACAAACTTTGCAGGTCCTTTAAATAATTTT 420
                ************************************************************

wtDCC43         ATTTTAGGAATTATAGCTTTTATCATTATCACTTTTATGCAAGGTGGTGTACCTTCAAAT 480
G1              ATTTTAGGAATTATAGCTTTTATCATTATCACTTTTATGCAAGGTGGTGTACCTTCAAAT 480
G2              ATTTTAGGAATTATAGCTTTTATCATTATCACTTTTATGTAAGGTGGTGTACCTTCAAAT 480
G3              ATTTTAGGAATTATAGCTTTTATCATTATCACTTTTATGCAAGGTGGTGTACCTTCAAAT 480
G4              ATTTTAGGAATTATAGCTTTTATCATTATCACTTTTATGCAAGGTGGTGTACCTTCAAAT 480
G5              ATTTTAGGAATTATAGCTTTTATCATTATCACTTTTATGCAAGGTGGTGTACCTTCAAAT 480
G6              ATTTTAGGAATTATAGCTTTTATCATTATCACTTTTATGCAAGGTGGTGTACCTTCAAAT 480
G7              ATTTTAGGAATTATAGCTTTTATCATTATCACTTTTATGCAAGGTGGTGTACCTTCAAAT 480
G8              ATTTTAGGAATTATAGCTTTTATCATTATCACTTTTATGCAAGGTGGTGTACCTTCAAAT 480
G9              ATTTTAGGAATTATAGCTTTTATCATTATCACTTTTATGCAAGGTGGTGTACCTTCAAAT 480
G10             ATTTTAGGAATTATAGCTTTTATCATTATCACTTTTATGCAAGGTGGTGTACCTTCAAAT 480
                ************************************ ********************

wtDCC43         ACAAATCAGATTGGACAAATTCAATCAGGGACACCAGCTTATACGGCAGGGCTGAAACCT 540
G1              ACAAATCAGATTGGACAAATTCAATCAGGGACACCAGCTTATACGGCAGGGCTGAAACCT 540
G2              ACAAATCAGATTGGACAAATTCAATCAGGGACACCAGCTTATACGGCAGGGCTGAAACCT 540
G3              ACAAATCAGATTGGACAAATTCAATCAGGGACACCAGCTTATACGGCAGGGCTGAAACCT 540
G4              ACAAATCAGATTGGACAAATTCAATCAGGGACACCAGCTTATACGGCAGGGCTGAAACCT 540
G5              ACAAATCAGATTGGACAAATTCAATCAGGGACACCAGCTTATACGGCAGGGCTGAAACCT 540
G6              ACAAATCAGATTGGACAAATTTAATCAGGGACACCAGCTTATACGGCAGGGCTGAAACCT 540
G7              ACAAATCAGATTGGACAAATTCAATCAGGGACACCAGCTTATACGGCAGGGCTGAAACCT 540
G8              ACAAATCAGATTGGACAAATTCAATCAGGGACACCAGCTTATACGGCAGGGCTGAAACCT 540
G9              ACAAATCAGATTGGACAAATTCAATCAGGGACACCAGCTTATACGGCAGGGCTGAAACCT 540
G10             ACAAATCAGATTGGACAAATTCAATCAGGGACACCAGCTTATACGGCAGGGCTGAAACCT 540
                ******************** ************************************

wtDCC43         AAGGACAAAATCCAAGCCGTAGATGGACATGAAACCAACTCCTGGAATGAAGTTGTTCAA 600
G1              AAGGACAAAATCCAAGCCGTAGATGGACATGAAACCAACTCCTGGAATGAAGTTGTTCAA 600
G2              AAGGACAAAATCCAAGCCGTAGATGGACATGAAACCAACTCCTGGAATGAAGTTGTTCAA 600
G3              AAGGACAAAATCCAAGCCGTAGATGGACATGAAACCAACTCCTGGAATGAAGTTGTTCAA 600
G4              AAGGACAAAATCCAAGCCGTAGATGGACATGAAACCAACTCCTGGAATGAAGTTGTTCAA 600
G5              AAGGACAAAATCCAAGCCGTAGATGGACATGAAACCAACTCCTGGAATGAAGTTGTTCAA 600
G6              AAGGACAAAATCCAAGCCGTAGATGGACATGAAACCAACTCCTGGAATGAAGTTGTTCAA 600
G7              AAGGACAAAATCCAAGCCGTAGATGGACATGAAACCAACTCCTGGAATGAAGTTGTTCAA 600
G8              AAGGACAAAATCCAAGCCGTAGATGGACATGAAACCAACTCCTGGAATGAAGTTGTTCAA 600
G9              AAGGACAAAATCCAAGCCGTAGATGGACATGAAACCAACTCCTGGAATGAAGTTGTTCAA 600
G10             AAGGACAAAATCCAAGCCGTAGATGGACATGAAACCAACTCCTGGAATGAAGTTGTTCAA 600
                ************************************************************

wtDCC43         CGCATAAGTGCAAATGAAGGAAAAGCCATCACGTTTGATATCGAACGGGATGGGACCAAC 660
G1              CGCATAAGTGCAAATGAAGGAAAAGCCATCACGTTTGATATCGAACGGGATGGGACCAAC 660
G2              CGCATAAGTGCAAATGAAGGAAAAGCCATCACGTTTGATATCGAACGGGATGGGACCAAC 660
G3              CGCATAAGTGCAAATGAAGGAAAAGCCATCACGTTTGATATCGAACGGGATGGGACCAAC 660
G4              CGCATAAGTGCAAATGAAGGAAAAGCCATCACGTTTGATATCGAACGGGATGGGACCAAC 660
G5              CGCATAAGTGCAAATGAAGGAAAAGCCATCACGTTTGATATCGAACGGGATGGGACCAAC 660
G6              CGCATAAGTGCAAATGAAGGAAAAGCCATCACGTTTGATATCGAACGGGATGGGACCAAC 660
G7              CGCATAAGTGCAAATGAAGGAAAAGCCATCACGTTTGATATCGAACGGGATGGGACCAAC 660
G8              CGCATAAGTGCAAATGAAGGAAAAGCCATCACGTTTGATATCGAACGGGATGGGACCAAC 660
G9              CGCATAAGTGCAAATGAAGGAAAAGCCATCACGTTTGATATCGAACGGGATGGGACCAAC 660
G10             CGCATAAGTGCAAATGAAGGAAAAGCCATCACGTTTGATATCGAACGGGATGGGACCAAC 660
                ************************************************************

wtDCC43         AAGTCAGTAGAAATTAAACCTGAAAAAATAGAAGGTAGATACCGTATTGGTATCACCCAA 720
G1              AAGTCAGTAGAAATTAAACCTGAAAAAATAGAAGGTAGATACCGTATTGGTATCACCCAA 720
G2              AAGTCAGTAGAAATTAAACCTGAAAAAATAGAAGGTAGATACCGTATTGGTATCACCCAA 720
G3              AAGTCAGTAGAAATTAAACCTGAAAAAATAGAAGGTAGATACCGTATTGGTATCACCCAA 720
G4              AAGTCAGTAGAAATTAAACCTGAAAAAATAGAAGGTAGATACCGTATTGGTATCACCCAA 720
G5              AAGTCAGTAGAAATTAAACCTGAAAAAATAGAAGGTAGATACCGTATTGGTATCACCTAA 720
G6              AAGTCAGTAGAAATTAAACCTGAAAAAATAGAAGGTAGATACCGTATTGGTATCACCCAA 720
G7              AAGTCAGTAGAAATTAAACCTGAAAAAATAGAAGGTAGATACCGTATTGGTATCACCCAA 720
G8              AAGTCAGTAGAAATTAAACCTGAAAAAATAGAAGGTAGATACCGTATTGGTATCACCCAA 720
G9              AAGTCAGTAGAAATTAAACCTGAAAAAATAGAAGGTAGATACCGTATTGGTATCACCCAA 720
G10             AAGTCAGTAGAAATTAAACCTGAAAAAATAGAAGGTAGATACCGTATTGGTATCACCCAA 720
                ********************************************************** **

wtDCC43         TCCATCAAGACGGGCTTTATAGATAAACTCACAGGTGGGTTTACACAGGCGATCAATGCC 780
G1              TCCATCAAGACGGGCTTTATAGATAAACTCACAGGTGGGTTTACACAGGCGATCAATGCC 780
G2              TCCATCAAGACGGGCTTTATAGATAAACTCACAGGTGGGTTTACACAGGCGATCAATGCC 780
G3              TCCATCAAGACGGGCTTTATAGATAAACTCACAGGTGGGTTTACACAGGCGATCAATGCC 780
G4              TCCATCAAGACGGGCTTTATAGATAAACTCACAGGTGGGTTTACACAGGCGATCAATGCC 780
```

```
G5              TCCATCAAGACGGGCTTTATAGATAAACTCACAGGTGGGTTTACACAGGCGATCAATGCC 780
G6              TCCATCAAGACGGGCTTTATAGATAAACTCACAGGTGGGTTTACACAGGCGATCAATGCC 780
G7              TCCATCAAGACGGGCTTTATAGATAAACTCACAGGTGGGTTTACACAGGCGATCAATGCC 780
G8              TCCATCAAGACGGGCTTTATAGATAAACTCACAGGTGGGTTTACACAGGCGATCAATGCC 780
G9              TCCATCAAGACGGGCTTTATAGATAAACTCACAGGTGGGTTTACACAGGCGATCAATGCC 780
G10             TCCATCAAGACGGGCTTTATAGATAAACTCACAGGTGGGTTTACACAGGCGATCAATGCC 780
                ************************************************************

wtDCC43         ACGACTCTAATCTTTAAAGCTTTAGGTAACCTCATCTCCCAACCGAGTCTTGACAAGCTG 840
G1              ACGACTCTAATCTTTAAAGCTTTAGGTAACCTCATCTCCCAACCGAGTCTTGACAAGCTG 840
G2              ACGACTCTAATCTTTAAAGCTTTAGGTAACCTCATCTCCCAACCGAGTCTTGACAAGCTG 840
G3              ACGACTCTAATCTTTAAAGCTTTAGGTAACCTCATCTCCCAACCGAGTCTTGACAAGCTG 840
G4              ACGACTCTAATCTTTAAAGCTTTAGGTAACCTCATCTCCCAACCGAGTCTTGACAAGCTG 840
G5              ACGACTCTAATCTTTAAAGCTTTAGGTAACCTCATCTCCCAACCGAGTCTTGACAAGCTG 840
G6              ACGACTCTAATCTTTAAAGCTTTAGGTAACCTCATCTCCCAACCGAGTCTTGACAAGCTG 840
G7              ACGACTCTAATCTTTAAAGCTTTAGGTAACCTCATCTCCCAACCGAGTCTTGACAAGCTG 840
G8              ACGACTCTAATCTTTAAAGCTTTAGGTAACCTCATCTCCCAACCGAGTCTTGACAAGCTG 840
G9              ACGACTCTAATCTTTAAAGCTTTAGGTAACCTCATCTCCCAACCGAGTCTTGACAAGCTG 840
G10             ACGACTCTAATCTTTAAAGCTTTAGGTAACCTCATCTCCCAACCGAGTCTTGACAAGCTG 840
                ************************************************************

wtDCC43         GGGGGG-CCGGTTGCTATCTATGACTTAAGTGGTCAGGCAGCGCGTGAAGGAATAGTATC 899
G1              GGGGG--CCGGTTGCTATCTATGACTTAAGTGGTCAGGCAGCGCGTGAAGGAATAGTATC 898
G2              GGGGGG-CCGGTTGCTATCTATGACTTAAGTGGTCAGGCAGCGCGTGAAGGAATAGTATC 899
G3              GGGGGGGCCCGGTTGCTATCTATGACTTAAGTGGTCAGGCAGCGCGTGAAGGAATAGTATC 900
G4              GGGGG--CCGGTTGCTATCTATGACTTAAGTGGTCAGGCAGCGCGTGAAGGAATAGTATC 898
G5              GGGGGG-CCGGTTGCTATCTATGACTTAAGTGGTCAGGCAGCGCGTGAAGGAATAGTATC 899
G6              GGGGGG-CCGGTTGCTATCTATGACTTAAGTGGTCAGGCAGCGCGTGAAGGAATAGTATC 899
G7              GGGGGG-CCGGTTGCTATCTATGACTTAAGTGGTCAGGCAGCGCGTGAAGGAATAGTATC 899
G8              GGGGGGGCCGGTTGCTATCTATGACTTAAGTGGTCAGGCAGCGCGTGAAGGAATAGTATC 900
G9              GGGGGGGCCGGTTGCTATCTATGACTTAAGTGGTCAGGCAGCGCGTGAAGGAATAGTATC 900
G10             GGGGG--CCGGTTGCTATCTATGACTTAAGTGGTCAGGCAGCGCGTGAAGGAATAGTATC 898
                *****  ************************************************** 

wtDCC43         TGTTATTGCTTTGCTTGCTATGCTGTCCATTAATCTCGGAATTGTAAATCTCATTCCAAT 959
G1              TGTTATTGCTTTGCTTGCTATGCTGTCCATTAATCTCGGAATTGTAAATCTCATTCCAAT 958
G2              TGTTATTGCTTTGCTTGCTATGCTGTCCATTAATCTCGGAATTGTAAATCTCATTCCAAT 959
G3              TGTTATTGCTTTGCTTGCTATGCTGTCCATTAATCTCGGAATTGTAAATCTCATTCCAAT 960
G4              TGTTATTGCTTTGCTTGCTATGCTGTCCATTAATCTCGGAATTGTAAATCTCATTCCAAT 958
G5              TGTTATTGCTTTGCTTGCTATGCTGTCCATTAATCTCGGAATTGTAAATCTCATTCCAAT 959
G6              TGTTATTGCTTTGCTTGCTATGCTGTCCATTAATCTCGGAATTGTAAATCTCATTCCAAT 959
G7              TGTTATTGCTTTGCTTGCTATGCTGTCCATTAATCTCGGAATTGTAAATCTCATTCCAAT 959
G8              TGTTATTGCTTTGCTTGCTATGCTGTCCATTAATCTCGGAATTGTAAATCTCATTCCAAT 960
G9              TGTTATTGCTTTGCTTGCTATGCTGTCCATTAATCTCGGAATTGTAAATCTCATTCCAAT 960
G10             TGTTATTGCTTTGCTTGCTATGCTGTCCATTAATCTCGGAATTGTAAATCTCATTCCAAT 958
                ************************************************************

wtDCC43         ACCTGTTCTTGATGGCGGTAAAATTGTTCTTAACATAGTCGAAGCGTTACGTGGTAAACC 1019
G1              ACCTGTTCTTGATGGCGGTAAAATTGTTCTTAACATAGTCGAAGCGTTACGTGGTAAACC 1018
G2              ACCTGTTCTTGATGGCGGTAAAATTGTTCTTAACATAGTCGAAGCGTTACGTGGTAAACC 1019
G3              ACCTGTTCTTGATGGCGGTAAAATTGTTCTTAACATAGTCGAAGCGTTACGTGGTAAACC 1020
G4              ACCTGTTCTTGATGGCGGTAAAATTGTTCTTAACATAGTCGAAGCGTTACGTGGTAAACC 1018
G5              ACCTGTTCTTGATGGCGGTAAAATTGTTCTTAACATAGTCGAAGCGTTACGTGGTAAACC 1019
G6              ACCTGTTCTTGATGGCGGTAAAATTGTTCTTAACATAGTCGAAGCGTTACGTGGTAAACC 1019
G7              ACCTGTTCTTGATGGCGGTAAAATTGTTCTTAACATAGTCGAAGCGTTACGTGGTAAACC 1019
G8              ACCTGTTCTTGATGGCGGTAAAATTGTTCTTAACATAGTCGAAGCGTTACGTGGTAAACC 1020
G9              ACCTGTTCTTGATGGCGGTAAAATTGTTCTTAACATAGTCGAAGCGTTACGTGGTAAACC 1020
G10             ACCTGTTCTTGATGGCGGTAAAATTGTTCTTAACATAGTCGAAGCGTTACGTGGTAAACC 1018
                ************************************************************

wtDCC43         ACTTTCACAAGAAAAAGAAGGAGTATTAACTCTTATTGGTCTTGTCTTTATGGTTGTCCT 1079
G1              ACTTTCACAAGAAAAAGAAGGAGTATTAACTCTTATTGGTCTTGTCTTTATGGTTGTCCT 1078
G2              ACTTTCACAAGAAAAAGAAGGAGTATTAACTCTTATTGGTCTTGTCTTTATGGTTGTCCT 1079
G3              ACTTTCACAAGAAAAAGAAGGAGTATTAACTCTTATTGGTCTTGTCTTTATGGTTGTCCT 1080
G4              ACTTTCACAAGAAAAAGAAGGAGTATTAACTCTTATTGGTCTTGTCTTTATGGTTGTCCT 1078
G5              ACTTTCACAAGAAAAAGAAGGAGTATTAACTCTTATTGGTCTTGTCTTTATGGTTGTCCT 1079
G6              ACTTTCACAAGAAAAAGAAGGAGTATTAACTCTTATTGGTCTTGTCTTTATGGTTGTCCT 1079
G7              ACTTTCACAAGAAAAAGAAGGAGTATTAACTCTTATTGGTCTTGTCTTTATGGTTGTCCT 1079
G8              ACTTTCACAAGAAAAAGAAGGAGTATTAACTCTTATTGGTCTTGTCTTTATGGTTGTCCT 1080
G9              ACTTTCACAAGAAAAAGAAGGAGTATTAACTCTTATTGGTCTTGTCTTTATGGTTGTCCT 1080
G10             ACTTTCACAAGAAAAAGAAGGAGTATTAACTCTTATTGGTCTTGTCTTTATGGTTGTCCT 1078
                ************************************************************

wtDCC43         CATGATCGCTGTGACTTGGAATGATATTCTCCGCGCTTTTATCAGATAA 1128
G1              CATGATCGCTGTGACTTGGAATGATATTCTCCGCGCTTTTATCAGATAA 1127
G2              CATGATCGCTGTGACTTGGAATGATATTCTCCGCGCTTTTATCAGATAA 1128
G3              CATGATCGCTGTGACTTGGAATGATATTCTCCGCGCTTTTATCAGATAA 1129
G4              CATGATCGCTGTGACTTGGAATGATATTCTCCGCGCTTTTATCAGATAA 1127
G5              CATGATCGCTGTGACTTGGAATGATATTCTCCGCGCTTTTATCAGATAA 1128
G6              CATGATCGCTGTGACTTGGAATGATATTCTCCGCGCTTTTATCAGATAA 1128
G7              CATGATCGCTGTGACTTGGAATGATATTCTCCGCGCTTTTATCAGATAA 1128
G8              CATGATCGCTGTGACTTGGAATGATATTCTCCGCGCTTTTATCAGATAA 1129
G9              CATGATCGCTGTGACTTGGAATGATATTCTCCGCGCTTTTATCAGATAA 1129
G10             CATGATCGCTGTGACTTGGAATGATATTCTCCGCGCTTTTATCAGATAA 1127
                ************************************************
```

A6: MSA of DCC43 RseP

```
wtDCC43         MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASVMIVDNQVVKINLS 60
G1              MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASVMIVDNQVVKINLS 60
G2              MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASVMIVDNQVVKINLS 60
```

```
G3              MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASVMIVDNQVVKINLS 60
G4              MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASVMIVDNQVVKINLS 60
G5              MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASVMIVDNQVVKINLS 60
G6              MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASVMIVDNQVVKINLS 60
G7              MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASVMIVDNQVVKINLS 60
G8              MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASVMIVDNQVVKINLS 60
G9              MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASVMIVDNQVVKINLS 60
G10             MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASVMIVDNQVVKINLS 60
                ************************************************************

wtDCC43         DRVTLENSVPMLVTEYDFEKELFIEGEVFGEITRYSVDHDATIIEGDGTEIRIAPLDVQY 120
G1              DRVTLENSVPMLVTEYDFEKELFIEGEVFGEITRYSVDHDATIIEGDGTEIRIAPLDVQY 120
G2              DRVTLENSVPMLVTEYDFEKELFIEGEVFGEITRYSVDHDATIIEGDGTEIRIAPLDVQY 120
G3              DRVTLENSVPMLVTEYDFEKELFIEGEVFGEITRYSVDHDATIIEGDGTEIRIAPLDVQY 120
G4              DRVTLENSVPMLVTEYDFEKELFIEGEVFGEITRYSVDHDATIIEGDGTEIRIAPLDVQY 120
G5              DRVTLENSVPMLVTEYDFEKELFIEGEVFGEITRYSVDHDATIIEGDGTEIRIAPLDVQY 120
G6              DRVTLENSVPMLVTEYDFEKELFIEGEVFGEITRYSVDHDATIIEGDGTEIRIAPLDVQY 120
G7              DRVTLENSVPMLVTEYDFEKELFIEGEVFGEITRYSVDHDATIIEGDGTEIRIAPLDV-- 118
G8              DRVTLENSVPMLVTEYDFEKELFIEGEVFGEITRYSVDHDATIIEGDGTEIRIAPLDVQY 120
G9              DRVTLENSVPMLVTEYDFEKELFIEGEVFGEITRYSVDHDATIIEGDGTEIRIAPLDVQY 120
G10             DRVTLENSVPMLVTEYDFEKELFIEGEVFGEITRYSVDHDATIIEGDGTEIRIAPLDVQY 120
                ************************************************************

wtDCC43         QSASVIGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQIQSGTPAYTAGLKP 180
G1              QSASVIGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQIQSGTPAYTAGLKP 180
G2              QSASVIGKILTNFAGPLNNFILGIIAFIIITFM---------------------------- 153
G3              QSASVIGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQIQSGTPAYTAGLKP 180
G4              QSASVIGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQIQSGTPAYTAGLKP 180
G5              QSASVIGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQIQSGTPAYTAGLKP 180
G6              QSASVIGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQI-------------- 167
G7              ------------------------------------------------------------
G8              QSASVIGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQIQSGTPAYTAGLKP 180
G9              QSASVIGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQIQSGTPAYTAGLKP 180
G10             QSASVIGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQIQSGTPAYTAGLKP 180


wtDCC43         KDKIQAVDGHETNSWNEVVQRISANEGKAITFDIERDGTNKSVEIKPEKIEGRYRIGITQ 240
G1              KDKIQAVDGHETNSWNEVVQRISANEGKAITFDIERDGTNKSVEIKPEKIEGRYRIGITQ 240
G2              ------------------------------------------------------------
G3              KDKIQAVDGHETNSWNEVVQRISANEGKAITFDIERDGTNKSVEIKPEKIEGRYRIGITQ 240
G4              KDKIQAVDGHETNSWNEVVQRISANEGKAITFDIERDGTNKSVEIKPEKIEGRYRIGITQ 240
G5              KDKIQAVDGHETNSWNEVVQRISANEGKAITFDIERDGTNKSVEIKPEKIEGRYRIGIT- 239
G6              ------------------------------------------------------------
G7              ------------------------------------------------------------
G8              KDKIQAVDGHETNSWNEVVQRISANEGKAITFDIERDGTNKSVEIKPEKIEGRYRIGITQ 240
G9              KDKIQAVDGHETNSWNEVVQRISANEGKAITFDIERDGTNKSVEIKPEKIEGRYRIGITQ 240
G10             KDKIQAVDGHETNSWNEVVQRISANEGKAITFDIERDGTNKSVEIKPEKIEGRYRIGITQ 240


wtDCC43         SIKTGFIDKLTGGFTQAINATTLIFKALGNLISQPSLDKLGGPVAIYDLSGQAAREGIVS 300
G1              SIKTGFIDKLTGGFTQAINATTLIFKALGNLISQPSLDKLGGRLLSMT------------ 288
G2              ------------------------------------------------------------
G3              SIKTGFIDKLTGGFTQAINATTLIFKALGNLISQPSLDKLGGAGCYL------------- 287
G4              SIKTGFIDKLTGGFTQAINATTLIFKALGNLISQPSLDKLGGRLLSMT------------ 288
G5              ------------------------------------------------------------
G6              ------------------------------------------------------------
G7              ------------------------------------------------------------
G8              SIKTGFIDKLTGGFTQAINATTLIFKALGNLISQPSLDKLGGAGCYL------------- 287
G9              SIKTGFIDKLTGGFTQAINATTLIFKALGNLISQPSLDKLGGAGCYL------------- 287
G10             SIKTGFIDKLTGGFTQAINATTLIFKALGNLISQPSLDKLGGRLLSMT------------ 288


wtDCC43         VIALLAMLSINLGIVNLIPIPVLDGGKIVLNIVEALRGKPLSQEKEGVLTLIGLVFMVVLMIAVTWNDILRAFIR 375
G1              ---------------------------------------------------------------------------
G2              ---------------------------------------------------------------------------
G3              ---------------------------------------------------------------------------
G4              ---------------------------------------------------------------------------
G5              ---------------------------------------------------------------------------
G6              ---------------------------------------------------------------------------
G7              ---------------------------------------------------------------------------
G8              ---------------------------------------------------------------------------
G9              ---------------------------------------------------------------------------
G10             ---------------------------------------------------------------------------
```

Lactococcus garvieae 1546

A7: MSA of 1546 *rseP*

```
wt1546          ATGGGCCCCAAAATTTTTGCACATCAAGCCAAAGATGGCACGTTTTATAGCATTCGTATA 60
G11             ATGGGCCCCAAAATTTTTGCACATCAAGCCAAAGATGGCACGTTTTATAGCATTCGTATA 60
G12             ATGGGCCCCAAAATTTTTGCACATCAAGCCAAAGATGGCACGTTTTATAGCATTCGTATA 60
G13             ATGGGCCCCAAAATTTTTGCACATCAAGCCAAAGATGGCACGTTTTATAGCATTCGTATA 60
G14             ATGGGCCCCAAAATTTTTGCACATCAAGCCAAAGATGGCACGTTTTATAGCATTCGTATA 60
G15             ATGGGCCCCAAAATTTTTGCACATCAAGCCAAAGATGGCACGTTTTATAGCATTCGTATA 60
G16             ATGGGCCCCAAAATTTTTGCACATCAAGCCAAAGATGGCACGTTTTATAGCATTCGTATA 60
G17             ATGGGCCCCAAAATTTTTGCACATCAAGCCAAAGATGGCACGTTTTATAGCATTCGTATA 60
G18             ATGGGCCCCAAAATTTTTGCACATCAAGCCAAAGATGGCACGTTTTATAGCATTCGTATA 60
G19             ATGGGCCCCAAAATTTTTGCACATCAAGCCAAAGATGGCACGTTTTATAGCATTCGTATA 60
G20             ATGGGCCCCAAAATTTTTGCACATCAAGCCAAAGATGGCACGTTTTATAGCATTCGTATA 60
                ************************************************************

wt1546          CTTCCTCTTGGAGGTTATGTCCGCATGGCGGGCTGGGGCGATGATACAACGGAAATAAAA 120
G11             CTTCCTCTTGGAGGTTATGTCCGCATGGCGGGCTGGGGCGATGATACAACGGAAATAAAA 120
G12             CTTCCTCTTGGAGGTTATGTCCGCATGGCGGGCTGGGGCGATGATACAACGGAAATAAAA 120
G13             CTTCCTCTTGGAGGTTATGTCCGCATGGCGGGCTGGGGCGATGATACAACGGAAATAAAA 120
G14             CTTCCTCTTGGAGGTTATGTCCGCATGGCGGGCTGGGGCGATGATACAACGGAAATAAAA 120
G15             CTTCCTCTTGGAGGTTATGTCCGCATGGCGGGCTGGGGCGATGATACAACGGAAATAAAA 120
G16             CTTCCTCTTGGAGGTTATGTCCGCATGGCGGGCTGGGGCGATGATACAACGGAAATAAAA 120
G17             CTTCCTCTTGGAGGTTATGTCCGCATGGCGGGCTGGGGCGATGATACAACGGAAATAAAA 120
G18             CTTCCTCTTGGAGGTTATGTCCGCATGGCGGGCTGGGGCGATGATACAACGGAAATAAAA 120
G19             CTTCCTCTTGGAGGTTATGTCCGCATGGCGGGCTGGGGCGATGATACAACGGAAATAAAA 120
G20             CTTCCTCTTGGAGGTTATGTCCGCATGGCGGGCTGGGGCGATGATACAACGGAAATAAAA 120
                ************************************************************

wt1546          AAAGGAACCCCGGCATCTTTGATGATTGTTGAAAATCAGGTGGTTAAAATTAATCTTTCT 180
G11             AAAGGAACCCCGGCATCTTTGATGATTGTTGAAAATCAGGTGGTTAAAATTAATCTTTCT 180
G12             AAAGGAACCCCGGCATCTTTGATGATTGTTGAAAATCAGGTGGTTAAAATTAATCTTTCT 180
G13             AAAGGAACCCCGGCATCTTTGATGATTGTTGAAAATCAGGTGGTTAAAATTAATCTTTCT 180
G14             AAAGGAACCCCGGCATCTTTGATGATTGTTGAAAATCAGGTGGTTAAAATTAATCTTTCT 180
G15             AAAGGAACCCCGGCATCTTTGATGATTGTTGAAAATCAGGTGGTTAAAATTAATCTTTCT 180
G16             AAAGGAACCCCGGCATCTTTGATGATTGTTGAAAATCAGGTGGTTAAAATTAATCTTTCT 180
G17             AAAGGAACCCCGGCATCTTTGATGATTGTTGAAAATCAGGTGGTTAAAATTAATCTTTCT 180
G18             AAAGGAACCCCGGCATCTTTGATGATTGTTGAAAATCAGGTGGTTAAAATTAATCTTTCT 180
G19             AAAGGAACCCCGGCATCTTTGATGATTGTTGAAAATCAGGTGGTTAAAATTAATCTTTCT 180
G20             AAAGGAACCCCGGCATCTTTGATGATTGTTGAAAATCAGGTGGTTAAAATTAATCTTTCT 180
                ************************************************************

wt1546          GAGCGGGTGACTTTGGAAAATAGTGTCCCCATGCTGGTCACTGAATACGATTTTGAAGAA 240
G11             GAGCGGGTGACTTTGGAAAATAGTGTCCCCATGCTGGTCACTGAATACGATTTTGAAGAA 240
G12             GAGCGGGTGACTTTGGAAAATAGTGTCCCCATGCTGGTCACTGAATACGATTTTGAAGAA 240
G13             GAGCGGGTGACTTTGGAAAATAGTGTCCCCATGCTGGTCACTGAATACGATTTTGAAGAA 240
G14             GAGCGGGTGACTTTGGAAAATAGTGTCCCCATGCTGGTCACTGAATACGATTTTGAAGAA 240
G15             GAGCGGGTGACTTTGGAAAATAGTGTCCCCATGCTGGTCACTGAATACGATTTTGAAGAA 240
G16             GAGCGGGTGACTTTGGAAAATAGTGTCCCCATGCTGGTCACTGAATACGATTTTGAAGAA 240
G17             GAGCGGGTGACTTTGGAAAATAGTGTCCCCATGCTGGTCACTGAATACGATTTTGAAGAA 240
G18             GAGCGGGTGACTTTGGAAAATAGTGTCCCCATGCTGGTCACTGAATACGATTTTGAAGAA 240
G19             GAGCGGGTGACTTTGGAAAATAGTGTCCCCATGCTGGTCACTGAATACGATTTTGAAGAA 240
G20             GAGCGGGTGACTTTGGAAAATAGTGTCCCCATGCTGGTCACTGAATACGATTTTGAAGAA 240
                ************************************************************

wt1546          GCACTTTTCATCGAAGGAGAAGTTTTTGGTGAAACAAAACGTTATTCAGTTGACCATGAT 300
G11             GCACTTTTCATCGAAGGAGAAGTTTTTGGTGAAACAAAACGTTATTCAGTTGACCATGAT 300
G12             GCACTTTTCATCGAAGGAGAAGTTTTTGGTGAAACAAAACGTTATTCAGTTGACCATGAT 300
G13             GCACTTTTCATCGAAGGAGAAGTTTTTGGTGAAACAAAACGTTATTCAGTTGACCATGAT 300
G14             GCACTTTTCATCGAAGGAGAAGTTTTTGGTGAAACAAAACGTTATTCAGTTGACCATGAT 300
G15             GCACTTTTCATCGAAGGAGAAGTTTTTGGTGAAACAAAACGTTATTCAGTTGACCATGAT 300
G16             GCACTTTTCATCGAAGGAGAAGTTTTTGGTGAAACAAAACGTTATTCAGTTGACCATGAT 300
G17             GCACTTTTCATCGAAGGAGAAGTTTTTGGTGAAACAAAACGTTATTCAGTTGACCATGAT 300
G18             GCACTTTTCATCGAAGGAGAAGTTTTTGGTGAAACAAAACGTTATTCAGTTGACCATGAT 300
G19             GCACTTTTCATCGAAGGAGAAGTTTTTGGTGAAACAAAACGTTATTCAGTTGACCATGAT 300
G20             GCACTTTTCATCGAAGGAGAAGTTTTTGGTGAAACAAAACGTTATTCAGTTGACCATGAT 300
                ************************************************************

wt1546          GCAACAATTATCGAAGCGGATGGCACGGAACTTCGTATCGCTCCACGCGACGTACAATAT 360
G11             GCAACAATTATCGAAGCGGATGGCACGGAACTTCGTATCGCTCCACGCGACGTACAATAT 360
G12             GCAACAATTATCGAAGCGGATGGCACGGAACTTCGTATCGCTCCACGCGACGTACAATAT 360
G13             GCAACAATTATCGAAGCGGATGGCACGGAACTTCGTATCGCTCCACGCGACGTACAATAT 360
G14             GCAACAATTATCGAAGCGGATGGCACGGAACTTCGTATCGCTCCACGCGACGTACAATAT 360
G15             GCAACAATTATCGAAGCGGATGGCACGGAACTTCGTATCGCTCCACGCGACGTACAATAT 360
G16             GCAACAATTATCGAAGCGGATGGCACGGAACTTCGTATCGCTCCACGCGACGTACAATAT 360
G17             GCAACAATTATCGAAGCGGATGGCACGGAACTTCGTATCGCTCCACGCGACGTACAATAT 360
G18             GCAACAATTATCGAAGCGGATGGCACGGAACTTCGTATCGCTCCACGCGACGTACAATAT 360
G19             GCAACAATTATCGAAGCGGATGGCACGGAACTTCGTATCGCTCCACGCGACGTACAATAT 360
G20             GCAACAATTATCGAAGCGGATGGCACGTAACTTCGTATCGCTCCACGCGACGTACAATAT 360
                ************************** ******************************
```

```
wt1546          CAAAGTGCCAGCGTTCTCGGCAAGATTTTAACAAATTTTGCTGGTCCTTTAAATAACTTT 420
G11             CAAAGTGCCAGCGTTCTCGGCAAGATTTTAACAAATTTTGCTGGTCCTTTAAATAACTTT 420
G12             CAAAGTGCCAGCGTTCTCGGCAAGATTTTAACAAATTTTGCTGGTCCTTTAAATAACTTT 420
G13             CAAAGTGCCAGCGTTCTCGGCAAGATTTTAACAAATTTTGCTGGTCCTTTAAATAACTTT 420
G14             CAAAGTGCCAGCGTTCTCGGCAAGATTTTAACAAATTTTGCTGGTCCTTTAAATAACTTT 420
G15             CAAAGTGCCAGCGTTCTCGGCAAGATTTTAACAAATTTTGCTGGTCCTTTAAATAACTTT 420
G16             CAAAGTGCCAGCGTTCTCGGCAAGATTTTAACAAATTTTGCTGGTCCTTTAAATAACTTT 420
G17             CAAAGTGCCAGCGTTCTCGGCAAGATTTTAACAAATTTTGCTGGTCCTTTAAATAACTTT 420
G18             CAAAGTGCCAGCGTTCTCGGCAAGATTTTAACAAATTTTGCTGGTCCTTTAAATAACTTT 420
G19             CAAAGTGCCAGCGTTCTCGGCAAGATTTTAACAAATTTTGCTGGTCCTTTAAATAACTTT 420
G20             CAAAGTGCCAGCGTTCTCGGCAAGATTTTAACAAATTTTGCTGGTCCTTTAAATAACTTT 420
                ************************************************************

wt1546          ATTTTAGGGATTATAGCTTTTATTATCATCACTTTTATGCAAGGTGGTGTACCTTCTAAT 480
G11             ATTTTAGGGATTATAGCTTTTATTATCATCACTTTTATGCAAGGTGGTGTACCTTCTAAT 480
G12             ATTTTAGGGATTATAGCTTTTATTATCATCACTTTTATGCAAGGTGGTGTACCTTCTAAT 480
G13             ATTTTAGGGATTATAGCTTTTATTATCATCACTTTTATGCAAGGTGGTGTACCTTCTAAT 480
G14             ATTTTAGGGATTATAGCTTTTATTATCATCACTTTTATGCAAGGTGGTGTACCTTCTAAT 480
G15             ATTTTAGGGATTATAGCTTTTATTATCATCACTTTTATGCAAGGTGGTGTACCTTCTAAT 480
G16             ATTTTAGGGATTATAGCTTTTATTATCATCACTTTTATGCAAGGTGGTGTACCTTCTAAT 480
G17             ATTTTAGGGATTATAGCTTTTATTATCATCACTTTTATGCAAGGTGGTGTACCTTCTAAT 480
G18             ATTTTAGGGATTATAGCTTTTATTATCATCACTTTTATGCAAGGTGGTGTACCTTCTAAT 480
G19             ATTTTAGGGATTATAGCTTTTATTATCATCACTTTTATGCAAGGTGGTGTACCTTCTAAT 480
G20             ATTTTAGGGATTATAGCTTTTATTATCATCACTTTTATGCAAGGTGGTGTACCTTCTAAT 480
                ************************************************************

wt1546          ACAAATCAAATCGGACAAGTCCAATCTGGCACCCCTGCTTATTCGGCAGGTTTACAGTCA 540
G11             ACAAATCAAATCGGACAAGTCCAATCTGGCACCCCTGCTTATTCGGCAGGTTTACAGTCA 540
G12             ACAAATCAAATCGGACAAGTCCAATCTGGCACCCCTGCTTATTCGGCAGGTTTACAGTCA 540
G13             ACAAATCAAATCGGACAAGTCCAATCTGGCACCCCTGCTTATTCGGCAGGTTTACAGTCA 540
G14             ACAAATCAAATCGGACAAGTCCAATCTGGCACCCCTGCTTATTCGGCAGGTTTACAGTCA 540
G15             ACAAATCAAATCGGACAAGTCCAATCTGGCACCCCTGCTTATTCGGCAGGTTTACAGTCA 540
G16             ACAAATCAAATCGGACAAGTCCAATCTGGCACCCCTGCTTATTCGGCAGGTTTACAGTCA 540
G17             ACAAATCAAATCGGACAAGTCCAATCTGGCACCCCTGCTTATTCGGCAGGTTTACAGTCA 540
G18             ACAAATCAAATCGGACAAGTCCAATCTGGCACCCCTGCTTATTCGGCAGGTTTACAGTCA 540
G19             ACAAATCAAATCGGACAAGTCCAATCTGGCACCCCTGCTTATTCGGCAGGTTTACAGTCA 540
G20             ACAAATCAAATCGGACAAGTCCAATCTGGCACCCCTGCTTATTCGGCAGGTTTACAGTCA 540
                ************************************************************

wt1546          AAGGATAAAATCCAAGCAGTCAATGGTGAAAAAACCAATAACTGGGAAGAAGTTGTCCAA 600
G11             AAGGATAAAATCCAAGCAGTCAATGGTGAAAAAACCAATAACTGGGAAGAAGTTGTCCAA 600
G12             AAGGATAAAATCCAAGCAGTCAATGGTGAAAAAACCAATAACTGGGAAGAAGTTGTCCAA 600
G13             AAGGATAAAATCCAAGCAGTCAATGGTGAAAAAACCAATAACTGGGAAGAAGTTGTCCAA 600
G14             AAGGATAAAATCCAAGCAGTCAATGGTGAAAAAACCAATAACTGGGAAGAAGTTGTCCAA 600
G15             AAGGATAAAATCCAAGCAGTCAATGGTGAAAAAACCAATAACTGGGAAGAAGTTGTCCAA 600
G16             AAGGATAAAATCCAAGCAGTCAATGGTGAAAAAACCAATAACTGGGAAGAAGTTGTCCAA 600
G17             AAGGATAAAATCCAAGCAGTCAATGGTGAAAAAACCAATAACTAGGAAGAAGTTGTCCAA 600
G18             AAGGATAAAATCCAAGCAGTCAATGGTGAAAAAACCAATAACTGGGAAGAAGTTGTCCAA 600
G19             AAGGATAAAATCCAAGCAGTCAATGGTGAAAAAACCAATAACTGGGAAGAAGTTGTCCAA 600
G20             AAGGATAAAATCCAAGCAGTCAATGGTGAAAAAACCAATAACTGGGAAGAAGTTGTCCAA 600
                ******************************************.*****************

wt1546          AGGATCGGTGCAAATGATGGCAGTGAAATCAAGCTGGATATCGAACGTGCAGGTGCGACG 660
G11             AGGATCGGTGCAAATGATGGCAGTGAAATCAAGCTGGATATCGAACGTGCAGGTGCGACG 660
G12             AGGATCGGTGCAAATGATGGCAGTGAAATCAAGCTGGATATCGAACGTGCAGGTGCGACG 660
G13             AGGATCGGTGCAAATGATGGCAGTGAAATCAAGCTGGATATCGAACGTGCAGGTGCGACG 660
G14             AGGATCGGTGCAAATGATGGCAGTGAAATCAAGCTGGATATCGAACGTGCAGGTGCGACG 660
G15             AGGATCGGTGCAAATGATGGCAGTGAAATCAAGCTGGATATCGAACGTGCAGGTGCGACG 660
G16             AGGATCGGTGCAAATGATGGCAGTGAAATCAAGCTGGATATCGAACGTGCAGGTGCGACG 660
G17             AGGATCGGTGCAAATGATGGCAGTGAAATCAAGCTGGATATCGAACGTGCAGGTGCGACG 660
G18             AGGATCGGTGCAAATGATGGCAGTGAAATCAAGCTGGATATCGAACGTGCAGGTGCGACG 660
G19             AGGATCGGTGCAAATGATGGCAGTGAAATCAAGCTGGATATCGAACGTGCAGGTGCGACG 660
G20             AGGATCGGTGCAAATGATGGCAGTGAAATCAAGCTGGATATCGAACGTGCAGGTGCGACG 660
                ************************************************************

wt1546          AAGACACTGGAGCTCAAGCCAGAAAAAATTGACGGAAGATACCGCATTGGTATCACTCAA 720
G11             AAGACACTGGAGCTCAAGCCAGAAAAAATTGACGGAAGATACCGCATTGGTATCACTCAA 720
G12             AAGACACTGGAGCTCAAGCCAGAAAAAATTGACGGAAGATACCGCATTGGTATCACTCAA 720
G13             AAGACACTGGAGCTCAAGCCAGAAAAAATTGACGGAAGATACCGCATTGGTATCACTCAA 720
G14             AAGACACTGGAGCTCAAGCCAGAAAAAATTGACGGAAGATACCGCATTGGTATCACTCAA 720
G15             AAGACACTGGAGCTCAAGCCAGAAAAAATTGACGGAAGATACCGCATTGGTATCACTCAA 720
G16             AAGACACTGGAGCTCAAGCCAGAAAAAATTGACGGAAGATACCGCATTGGTATCACTCAA 720
G17             AAGACACTGGAGCTCAAGCCAGAAAAAATTGACGGAAGATACCGCATTGGTATCACTCAA 720
G18             AAGACACTGGAGCTCAAGCCAGAAAAAATTGACGGAAGATACCGCATTGGTATCACTCAA 720
G19             AAGACACTGGAGCTCAAGCCAGAAAAAATTGACGGAAGATACCGCATTGGTATCACTCAA 720
G20             AAGACACTGGAGCTCAAGCCAGAAAAAATTGACGGAAGATACCGCATTGGTATCACTCAA 720
                ************************************************************

wt1546          ACAATCAAAACTGGCTTTTTAGATAAGCTAACGGGTGGCTTTACTCAGGCTGCTAACGCA 780
G11             ACAATCAAAACTGGCTTTTTAGATAAGCTAACGGGTGGCTTTACTCAGGCTGCTAACGCA 780
G12             ACAATCAAAACTGGCTTTTTAGATAAGCTAACGGGTGGCTTTACTCAGGCTGCTAACGCA 780
G13             ACAATCAAAACTGGCTTTTTAGATAAGCTAACGGGTGGCTTTACTCAGGCTGCTAACGCA 780
G14             ACAATCAAAACTGGCTTTTTAGATAAGCTAACGGGTGGCTTTACTCAGGCTGCTAACGCA 780
G15             ACAATCAAAACTGGCTTTTTAGATAAGCTAACGGGTGGCTTTACTCAGGCTGCTAACGCA 780
G16             ACAATCAAAACTGGCTTTTTAGATAAGCTAACGGGTGGCTTTACTCAGGCTGCTAACGCA 780
G17             ACAATCAAAACTGGCTTTTTAGATAAGCTAACGGGTGGCTTTACTCAGGCTGCTAACGCA 780
```

```
G18             ACAATCAAAACTGGCTTTTTAGATAAGCTAACGGGTGGCTTTACTCAGGCTGCTAACGCA 780
G19             ACAATCAAAACTGGCTTTTTAGATAAGCTAACGGGTGGCTTTACTCAGGCTGCTAACGCA 780
G20             ACAATCAAAACTGGCTTTTTAGATAAGCTAACGGGTGGCTTTACTCAGGCTGCTAACGCA 780
                ************************************************************

wt1546          ACAACGTTGATATTTAAAGCTTTAGGTAACCTCATTGCACAGCCAAGTCTTGATAAACTT 840
G11             ACAACGTTGATATTTAAAGCTTTAGGTAACCTCATTGCACAGCCAAGTCTTGATAAACTT 840
G12             ACAACGTTGATATTTAAAGCTTTAGGTAACCTCATTGCACAGCCAAGTCTTGATAAACTT 840
G13             ACAACGTTGATATTTAAAGCTTTAGGTAACCTCATTGCACAGCCAAGTCTTGATAAACTT 840
G14             ACAACGTTGATATTTAAAGCTTTAGGTAACCTCATTGCACAGCCAAGTCTTGATAAACTT 840
G15             ACAACGTTGATATTTAAAGCTTTAGGTAACCTCATTGCACAGCCAAGTCTTGATAAACTT 840
G16             ACAACGTTGATATTTAAAGCTTTAGGTAACCTCATTGCACAGCCAAGTCTTGATAAACTT 840
G17             ACAACGTTGATATTTAAAGCTTTAGGTAACCTCATTGCACAGCCAAGTCTTGATAAACTT 840
G18             ACAACGTTGATATTTAAAGCTTTAGGTAACCTCATTGCACAGCCAAGTCTTGATAAACTT 840
G19             ACAACGTTGATATTTAAAGCTTTAGGTAACCTCATTGCACAGCCAAGTCTTGATAAACTT 840
G20             ACAACGTTGATATTTAAAGCTTTAGGTAACCTCATTGCACAGCCAAGTCTTGATAAACTT 840
                ************************************************************

wt1546          GGTGGACCCGTTGCCATTTATGATTTAAGTGGGCAAGCAGCCCGAGAAGGTCTAGTATCT 900
G11             GGTGGACCCGTTGCCATTTATGATTTAAGTGGGCAAGCAGCCCGAGAAGGTCTAGTATCT 900
G12             GGTGGACCCGTTGCCATTTATGATTTAAGTGGGCAAGCAGCCCGAGAAGGTCTAGTATCT 900
G13             GGTGGACCCGTTGCCATTTATGATTTAAGTGGGCAAGCAGCCCGAGAAGGTCTAGTATCT 900
G14             GGTGGACCCGTTGCCATTTATGATTTAAGTGGGCAAGCAGCCCGAGAAGGTCTAGTATCT 900
G15             GGTGGACCCGTTGCCATTTATGATTTAAGTGGGCAAGCAGCCCGAGAAGGTCTAGTATCT 900
G16             GGTGGACCCGTTGCCATTTATGATTTAAGTGGGCAAGCAGCCCGAGAAGGTCTAGTATCT 900
G17             GGTGGACCCGTTGCCATTTATGATTTAAGTGGGCAAGCAGCCCGAGAAGGTCTAGTATCT 900
G18             GGTGGACCCGTTGCCATTTATGATTTAAGTGGGCAAGCAGCCCGAGAAGGTCTAGTATCT 900
G19             GGTGGACCCGTTGCCATTTATGATTTAAGTGGGCAAGCAGCCCGAGAAGGTCTAGTATCT 900
G20             GGTGGACCCGTTGCCATTTATGATTTAAGTGGGCAAGCAGCCCGAGAAGGTCTAGTATCT 900
                ************************************************************

wt1546          GTTATCGCTTTACTTGCCATGCTCTCAATTAATCTGGGAATCGTAAACTTAATCCCAATC 960
G11             GTTATCGCTTTACTTGCCATGCTCTCAATTAATCTGGGAATCGTAAACTTAATCCCAATC 960
G12             GTTATCGCTTTACTTGCCATGCTCTCAATTAATCTGGGAATCGTAAACTTAATCCCAATC 960
G13             GTTATCGCTTTACTTGCCATGCTCTCAATTAATCTGGGAATCGTAAACTTAATCCCAATC 960
G14             GTTATCGCTTTACTTGCCATGCTCTCAATTAATCTGGGAATCGTAAACTTAATCCCAATC 960
G15             GTTATCGCTTTACTTGCCATGCTCTCAATTAATCTGGGAATCGTAAACTTAATCCCAATC 960
G16             GTTATCGCTTTACTTGCCATGCTCTCAATTAATCTGGGAATCGTAAACTTAATCCCAATC 960
G17             GTTATCGCTTTACTTGCCATGCTCTCAATTAATCTGGGAATCGTAAACTTAATCCCAATC 960
G18             GTTATCGCTTTACTTGCCATGCTCTCAATTAATCTGGGAATCGTAAACTTAATCCCAATC 960
G19             GTTATCGCTTTACTTGCCATGCTCTCAATTAATCTGGGAATCGTAAACTTAATCCCAATC 960
G20             GTTATCGCTTTACTTGCCATGCTCTCAATTAATCTGGGAATCGTAAACTTAATCCCAATC 960
                ************************************************************

wt1546          CCCGTCCTTGACGGAGGGAAAATTGTCCTCAATATTGTTGAGGCCCTACGAGGCAAACCA 1020
G11             CCCGTCCTTGACGGAGGGAAAATTGTCCTCAATATTGTTGAGGCCCTACGAGGCAAACCA 1020
G12             CCCGTCCTTGACGGAGGGAAAATTGTCCTCAATATTGTTGAGGCCCTACGAGGCAAACCA 1020
G13             CCCGTCCTTGACGGAGGGAAAATTGTCCTCAATATTGTTGAGGCCCTACGAGGCAAACCA 1020
G14             CCCGTCCTTGACGGAGGGAAAATTGTCCTCAATATTGTTGAGGCCCTACGAGGCAAACCA 1020
G15             CCCGTCCTTGACGGAGGGAAAATTGTCCTCAATATTGTTGAGGCCCTACGAGGCAAACCA 1020
G16             CCCGTCCTTGACGGAGGGAAAATTGTCCTCAATATTGTTGAGGCCCTACGAGGCAAACCA 1020
G17             CCCGTCCTTGACGGAGGGAAAATTGTCCTCAATATTGTTGAGGCCCTACGAGGCAAACCA 1020
G18             CCCGTCCTTGACGGAGGGAAAATTGTCCTCAATATTGTTGAGGCCCTACGAGGCAAACCA 1020
G19             CCCGTCCTTGACGGAGGGAAAATTGTCCTCAATATTGTTGAGGCCCTACGAGGCAAACCA 1020
G20             CCCGTCCTTGACGGAGGGAAAATTGTCCTCAATATTGTTGAGGCCCTACGAGGCAAACCA 1020
                ************************************************************

wt1546          CTTTCACAAGAAAAAGAAGGAATCTTAACACTTGTTGGTCTTGTCTTTATGGTGGTCTTG 1080
G11             CTTTCACAAGAAAAAGAAGGAATCTTAACACTTGTTGGTCTTGTCTTTATGGTGGTCTTG 1080
G12             CTTTCACAAGAAAAAGAAGGAATCTTAACACTTGTTGGTCTTGTCTTTATGGTGGTCTTG 1080
G13             CTTTCACAAGAAAAAGAAGGAATCTTAACACTTGTTGGTCTTGTCTTTATGGTGGTCTTG 1080
G14             CTTTCACAAGAAAAAGAAGGAATCTTAACACTTGTTGGTCTTGTCTTTATGGTGGTCTTG 1080
G15             CTTTCACAAGAAAAAGAAGGAATCTTAACACTTGTTGGTCTTGTCTTTATGGTGGTCTTG 1080
G16             CTTTCACAAGAAAAAGAAGGAATCTTAACACTTGTTGGTCTTGTCTTTATGGTGGTCTTG 1080
G17             CTTTCACAAGAAAAAGAAGGAATCTTAACACTTGTTGGTCTTGTCTTTATGGTGGTCTTG 1080
G18             CTTTCACAAGAAAAAGAAGGAATCTTAACACTTGTTGGTCTTGTCTTTATGGTGGTCTTG 1080
G19             CTTTCACAAGAAAAAGAAGGAATCTTAACACTTGTTGGTCTTGTCTTTATGGTGGTCTTG 1080
G20             CTTTCACAAGAAAAAGAAGGAATCTTAACACTTGTTGGTCTTGTCTTTATGGTGGTCTTG 1080
                ************************************************************

wt1546          ATGATTGCTGTGACTTGGAATGATATCCTCCGAGCCTTTATTAGATAA 1128
G11             ATGATTGCTGTGACTTGGAATGATATCCTCCGAGCCTTTATTAGATAA 1128
G12             ATGATTGCTGTGACTTGGAATGATATCCTCCGAGCCTTTATTAGATAA 1128
G13             ATGATTGCTGTGACTTGGAATGATATCCTCCGAGCCTTTATTAGATAA 1128
G14             ATGATTGCTGTGACTTGGAATGATATCCTCCGAGCCTTTATTAGATAA 1128
G15             ATGATTGCTGTGACTTGGAATGATATCCTCCGAGCCTTTATTAGATAA 1128
G16             ATGATTGCTGTGACTTGGAATGATATCCTCCGAGCCTTTATTAGATAA 1128
G17             ATGATTGCTGTGACTTGGAATGATATCCTCCGAGCCTTTATTAGATAA 1128
G18             ATGATTGCTGTGACTTGGAATGATATCCTCCGAGCCTTTATTAGATAA 1128
G19             ATGATTGCTGTGACTTGGAATGATATCCTCCGAGCCTTTATTAGATAA 1128
G20             ATGATTGCTGTGACTTGGAATGATATCCTCCGAGCCTTTATTAGATAA 1128
                ************************************************
```

A8: MSA of 1546 RseP

```
wt1546          MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASLMIVENQVVKINLS 60
G11             MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASLMIVENQVVKINLS 60
G12             MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASLMIVENQVVKINLS 60
G13             MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASLMIVENQVVKINLS 60
G14             MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASLMIVENQVVKINLS 60
G15             MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASLMIVENQVVKINLS 60
G16             MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASLMIVENQVVKINLS 60
G17             MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASLMIVENQVVKINLS 60
G18             MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASLMIVENQVVKINLS 60
G19             MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASLMIVENQVVKINLS 60
G20             MGPKIFAHQAKDGTFYSIRILPLGGYVRMAGWGDDTTEIKKGTPASLMIVENQVVKINLS 60
                ************************************************************

wt1546          ERVTLENSVPMLVTEYDFEEALFIEGEVFGETKRYSVDHDATIIEADGTELRIAPRDVQY 120
G11             ERVTLENSVPMLVTEYDFEEALFIEGEVFGETKRYSVDHDATIIEADGTELRIAPRDVQY 120
G12             ERVTLENSVPMLVTEYDFEEALFIEGEVFGETKRYSVDHDATIIEADGTELRIAPRDVQY 120
G13             ERVTLENSVPMLVTEYDFEEALFIEGEVFGETKRYSVDHDATIIEADGTELRIAPRDVQY 120
G14             ERVTLENSVPMLVTEYDFEEALFIEGEVFGETKRYSVDHDATIIEADGTELRIAPRDVQY 120
G15             ERVTLENSVPMLVTEYDFEEALFIEGEVFGETKRYSVDHDATIIEADGTELRIAPRDVQY 120
G16             ERVTLENSVPMLVTEYDFEEALFIEGEVFGETKRYSVDHDATIIEADGTELRIAPRDVQY 120
G17             ERVTLENSVPMLVTEYDFEEALFIEGEVFGETKRYSVDHDATIIEADGTELRIAPRDVQY 120
G18             ERVTLENSVPMLVTEYDFEEALFIEGEVFGETKRYSVDHDATIIEADGTELRIAPRDVQY 120
G19             ERVTLENSVPMLVTEYDFEEALFIEGEVFGETKRYSVDHDATIIEADGTELRIAPRDVQY 120
G20             ERVTLENSVPMLVTEYDFEEALFIEGEVFGETKRYSVDHDATIIEADGT----------- 109
                *********************************************

wt1546          QSASVLGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQVQSGTPAYSAGLQS 180
G11             QSASVLGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQVQSGTPAYSAGLQS 180
G12             QSASVLGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQVQSGTPAYSAGLQS 180
G13             QSASVLGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQVQSGTPAYSAGLQS 180
G14             QSASVLGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQVQSGTPAYSAGLQS 180
G15             QSASVLGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQVQSGTPAYSAGLQS 180
G16             QSASVLGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQVQSGTPAYSAGLQS 180
G17             QSASVLGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQVQSGTPAYSAGLQS 180
G18             QSASVLGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQVQSGTPAYSAGLQS 180
G19             QSASVLGKILTNFAGPLNNFILGIIAFIIITFMQGGVPSNTNQIGQVQSGTPAYSAGLQS 180
G20             ------------------------------------------------------------ 

wt1546          KDKIQAVNGEKTNNWEEVVQRIGANDGSEIKLDIERAGATKTLELKPEKIDGRYRIGITQ 240
G11             KDKIQAVNGEKTNNWEEVVQRIGANDGSEIKLDIERAGATKTLELKPEKIDGRYRIGITQ 240
G12             KDKIQAVNGEKTNNWEEVVQRIGANDGSEIKLDIERAGATKTLELKPEKIDGRYRIGITQ 240
G13             KDKIQAVNGEKTNNWEEVVQRIGANDGSEIKLDIERAGATKTLELKPEKIDGRYRIGITQ 240
G14             KDKIQAVNGEKTNNWEEVVQRIGANDGSEIKLDIERAGATKTLELKPEKIDGRYRIGITQ 240
G15             KDKIQAVNGEKTNNWEEVVQRIGANDGSEIKLDIERAGATKTLELKPEKIDGRYRIGITQ 240
G16             KDKIQAVNGEKTNNWEEVVQRIGANDGSEIKLDIERAGATKTLELKPEKIDGRYRIGITQ 240
G17             KDKIQAVNGEKTNN---------------------------------------------- 194
G18             KDKIQAVNGEKTNNWEEVVQRIGANDGSEIKLDIERAGATKTLELKPEKIDGRYRIGITQ 240
G19             KDKIQAVNGEKTNNWEEVVQRIGANDGSEIKLDIERAGATKTLELKPEKIDGRYRIGITQ 240
G20             ------------------------------------------------------------ 

wt1546          TIKTGFLDKLTGGFTQAANATTLIFKALGNLIAQPSLDKLGGPVAIYDLSGQAAREGLVS 300
G11             TIKTGFLDKLTGGFTQAANATTLIFKALGNLIAQPSLDKLGGPVAIYDLSGQAAREGLVS 300
G12             TIKTGFLDKLTGGFTQAANATTLIFKALGNLIAQPSLDKLGGPVAIYDLSGQAAREGLVS 300
G13             TIKTGFLDKLTGGFTQAANATTLIFKALGNLIAQPSLDKLGGPVAIYDLSGQAAREGLVS 300
G14             TIKTGFLDKLTGGFTQAANATTLIFKALGNLIAQPSLDKLGGPVAIYDLSGQAAREGLVS 300
G15             TIKTGFLDKLTGGFTQAANATTLIFKALGNLIAQPSLDKLGGPVAIYDLSGQAAREGLVS 300
G16             TIKTGFLDKLTGGFTQAANATTLIFKALGNLIAQPSLDKLGGPVAIYDLSGQAAREGLVS 300
G17             ------------------------------------------------------------ 
G18             TIKTGFLDKLTGGFTQAANATTLIFKALGNLIAQPSLDKLGGPVAIYDLSGQAAREGLVS 300
G19             TIKTGFLDKLTGGFTQAANATTLIFKALGNLIAQPSLDKLGGPVAIYDLSGQAAREGLVS 300
G20             ------------------------------------------------------------ 

wt1546          VIALLAMLSINLGIVNLIPIPVLDGGKIVLNIVEALRGKPLSQEKEGILTLVGLVFMVVLMIAVTWNDILRAFIR 375
G11             VIALLAMLSINLGIVNLIPIPVLDGGKIVLNIVEALRGKPLSQEKEGILTLVGLVFMVVLMIAVTWNDILRAFIR 375
G12             VIALLAMLSINLGIVNLIPIPVLDGGKIVLNIVEALRGKPLSQEKEGILTLVGLVFMVVLMIAVTWNDILRAFIR 375
G13             VIALLAMLSINLGIVNLIPIPVLDGGKIVLNIVEALRGKPLSQEKEGILTLVGLVFMVVLMIAVTWNDILRAFIR 375
G14             VIALLAMLSINLGIVNLIPIPVLDGGKIVLNIVEALRGKPLSQEKEGILTLVGLVFMVVLMIAVTWNDILRAFIR 375
G15             VIALLAMLSINLGIVNLIPIPVLDGGKIVLNIVEALRGKPLSQEKEGILTLVGLVFMVVLMIAVTWNDILRAFIR 375
G16             VIALLAMLSINLGIVNLIPIPVLDGGKIVLNIVEALRGKPLSQEKEGILTLVGLVFMVVLMIAVTWNDILRAFIR 375
G17             ------------------------------------------------------------------------------ 
G18             VIALLAMLSINLGIVNLIPIPVLDGGKIVLNIVEALRGKPLSQEKEGILTLVGLVFMVVLMIAVTWNDILRAFIR 375
G19             VIALLAMLSINLGIVNLIPIPVLDGGKIVLNIVEALRGKPLSQEKEGILTLVGLVFMVVLMIAVTWNDILRAFIR 375
G20             ------------------------------------------------------------------------------ 
```

Enterococcus faecium 2787

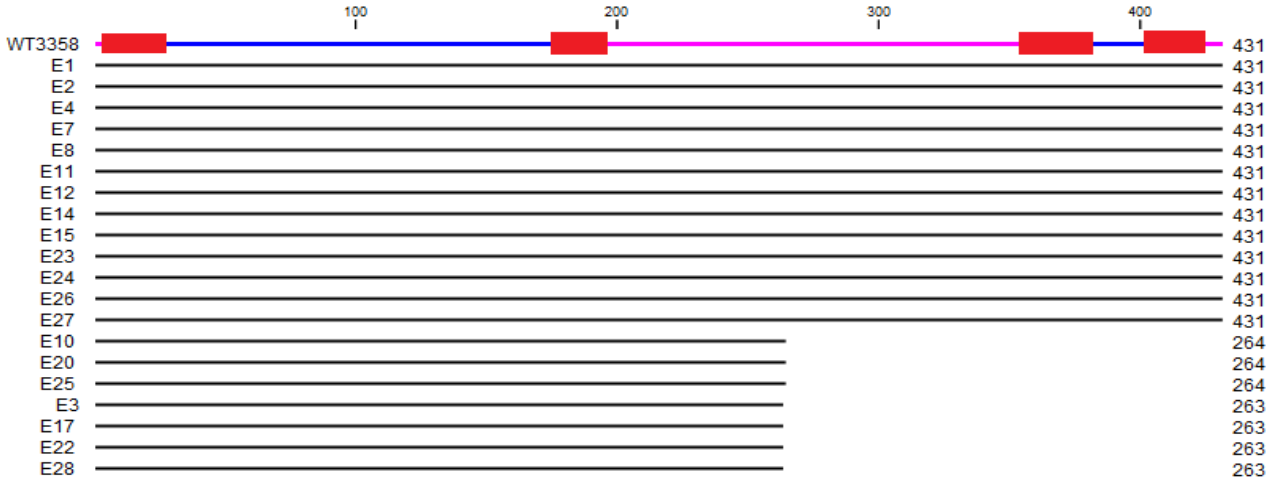
A9: MSA of 2787 *rseP* transposons

```
PO_B76          -----ATAAAGTCCATATAATTGTGTAAAAGATAAAAGGCCATGT---AACAGCCCTTTT 52
PO_B84          ----GATAAAGTCCATATAATTGTGTAAAAGATAAAAGGCCATGT---AACAGCCCTTTT 53
PO_B74          ----GATAAAGTCCATATAATTGTGTAAAAGATAAAAGGCCATGT---AACAGCCCTTTT 53
PO_B32          ----GGGAGCGTCAATAATTTTGTGTAAAT--AACTCGTCCTTCTGCAAAATAATGGGTT 54
PO_B80          ----GGGAGCGTCAATAATTTTGTGTAAAT--AACTCGTCCTTCTGCAAAATAATGGGTT 54
PO_B70          ----AGTCAAGTCCATAATCCTGTGTAAAA--TACTATACAATGT-----TTTACCGGTC 49
PO_B94          ----AGTCAAGTCCATAATCCTGTGTAAAA--TACTATACAATGT-----TTTACCGGTC 49
PO_B30          GTTTAGTCAAGTCCATAATCCTGTGTAAAA--TACTATACAATGT-----TTTACCGGTC 53
                     . ...***.***::  ********:  :*.:.  *.:* *      : .    *

PO_B76          ACGGTACAATGTTTTTAACCACAAAAACATACCCAGGAGG---ACGTTACATGACCCAAG 109
PO_B84          ACGGTACAATGTTTTTAACCACAAAAACATACCCAGGAGG---ACGTTACATGACCCAAG 110
PO_B74          ACGGTACAATGTTTTTAACCACAAAAACATACCCAGGAGG---ACGTTACATGACCCAAG 110
PO_B32          ACTCAGTAAACATTGAAGCTAATGTATCCGTCACTTGTTGGAAGCCTTTATGGCTTCTAT 114
PO_B80          ACTCAGTAAACATTGAAGCTAATGTATCCGTCACTTGTTGGAAGCCTTTATGGCTTCTAT 114
PO_B70          TCTCATTGATCA---AACTTAATATAT---TTTCTTGTAG---AACTG---AGCCATTCG 97
PO_B94          TCTCATTGATCA---AACTTAATATAT---TTTCTTGTAG---AACTG---AGCCATTCG 97
PO_B30          TCTCATTGATCA---AACTTAATATAT---TTTCTTGTAG---AACTG---AGCCATTCG 101
                :*  :  .*: :   :*   *.:.:*:   :  *: *: *   .. *     *.   :.

PO_B76          TACATTTTACACTGAACAACGAAGAGGTTCAAAGTATTATTGAACATTCGGTAAAAGATG 169
PO_B84          TACATTTTACACTGAACAACGAAGAGGTTCAAAGTATTATTGAACATTCGGTAAAAGATG 170
PO_B74          TACATTTTACACTGAACAACGAAGAGGTTCAAAGTATTATTGAACATTCGGTAAAAGATG 170
PO_B32          TTAAGAATTTTTGATTGTATGTATCAAAGATGGATACTAGGAAACGCTC--TAGTGATTC 172
PO_B80          TTAAGAATTTTTGATTGTATGTATCAAAGATGGATACTAGGAAACGCTC--TAGTGATTC 172
PO_B70          TCATGAAGGTCCATAAGGACAGCTCCAATTAATCGATTGGCAGACGTTCGGTTGGGAAAA 157
PO_B94          TCATGAAGGTCCATAAGGACAGCTCCAATTAATCGATTGGCAGACGTTCGGTTGGGAAAA 157
PO_B30          TCATGAAGGTCCATAAGGACAGCTCCAATTAATCGATTGGCAGACGTTCGGTTGGGAAAA 161
                * .: ::       ::  * . . . .:  :.   * *.  ..**. **  *:. ..::

PO_B76          ATGTTTCT-AAAAATATTTTAACCACTGTTTTCAACCAATTGATGGAAAATCAACGAACA 228
PO_B84          ATGTTTCT-AAAAATATTTTAACCACTGTTTTCAACCAATTGATGGAAAATCAACGAACA 229
PO_B74          ATGTTTCT-AAAAATATTTTAACCACTGTTTTCAACCAATTGATGGAAAATCAACGAACA 229
PO_B32          TTCATTTTGAAACTGATCTTTTCTGCGGCTGTATTTCTTGATTTGTTTATTGAAAGATTC 232
PO_B80          TTCATTTTGAAACTGCTCTTTTCTGCGGCTGTATTTCTTGATTTGTTTATTGAAAGATTC 232
PO_B70          ATCCGAAT-AATCTTTTCTCTTCTGCGGACTTCCTGGTTCAGTCGTTCAAG--AAGGTTG 214
PO_B94          ATCCGAAT-AATCTTTTCTCTTCTGCGGACTTCCTGGTTCAGTCGTTCAAG--AAGGTTG 214
PO_B30          ATCCGAAT-AATCTTTTCTCTTCTGCGGACTTCCTGGTTCAGTCGTTCAAG--AAGGTTG 218
                :*   : * **:.:  * * ::* .* *   *. :  :: : : * : *:   *.*.:

PO_B76          GAATATATTC-AAGCCGATGACTATGAACGTTCAGAAAGTCGTCAGAGTCAAAGAAATGG 287
PO_B84          GAATATATTC-AAGCCGATGACTATGAACGTTCAGAAAGTCGTCAGAGTCAAAGAAATGG 288
PO_B74          GAATATATTC-AAGCCGATGACTATGAACGTTCAGAAAGTCGTCAGAGTCAAAGAAATGG 288
PO_B32          GATTAAATTAGTTGAGTAAATGCTTCGGCGGATACCAGGCG--GAAACTCATAAAAAGT- 289
PO_B80          GATTAAATTAGTTGAGTAAATGCTTCGGCGGATACCAGGCG--GAAACTCATAAAAAGT- 289
PO_B70          GTGCTTTTTAGCCGATTGTGACCATTTCCGATAACCGTGTATTGAAAGGCATCTTCG--- 271
PO_B94          GTGCTTTTTAGCCGATTGTGACCATTTCCGATAACCGTGTATTGAAAGGCATCTTCG--- 271
PO_B30          GTGCTTTTTAGCCGATTGTGACCATTTCCGATAACCGTGTATTGAAAGGCATCTTCG--- 275
                *:  :::**.   *.  .:.:  :*   ** : * .. *     *.*  **:. :..

PO_B76          CTATTATGAGCGAGACTTTACGACTCGTGTGGGCACACTCGAATTAAAAGTGCCTAGAAC 347
PO_B84          CTATTATGAGCGAGACTTTACGACTCGTGTGGGCACACTCGAATTAAAAGTGCCTAGAAC 348
PO_B74          CTATTATGAGCGAGACTTTACGACTCGTGTGGGCACACTCGAATTAAAAGTGCCTAGAAC 348
PO_B32          CAATAAATCTTGATTTTTTATCAGTGACTGCGTCACTTTAGGAT--AAGTTTTCTG--CC 345
PO_B80          CAATAAATCTTGATTTTTTATCAGTGACTGCGTCACTTTAGGAT--AAGTTTTCTG--CC 345
PO_B70          AAGCCATTATCCAATGTTTCGCA-AGCTTTTGTATATTTAGACTGGTCGATATATT--CA 328
PO_B94          AAGCCATTATCCAATGTTTCGCA-AGCTTTTGTATATTTAGACTGGTCGATATATT--CA 328
PO_B30          AAGCCATTATCCAATGTTTCGCA-AGCTTTTGTATATTTAGACTGGTCGATATATT--CA 332
                .:.  *: .   * : ***.  * :      * .:.: *.*..*  :.. *  .*   ..

PO_B76          ACGTGATGGTGAATTTTCACCGACGGTGTTTGAGCGTTATCAGCGAAATGAAAAGGCACT 407
PO_B84          ACGTGATGGTGAATTT-CACCGACGGTGTTTGAGCGTTATCAGCGAAATGAAAAGGCACT 407
PO_B74          ACGTGATGGTGAATTTTCACCGACGGTGTTTGAGCGTTATCAGCGAAATGAAAAGGCACT 408
PO_B32          ACTTCTCAATCATACTCC-CTAAAAAGGTATTTGCCTCTTCCTTTGAGCTAGCTTGATAA 404
PO_B80          ACTTCTCAATCATACTCC-CTAAAAAGGTATTTGCCTCTTCCTTTGAGCTAGCTTGATAA 404
PO_B70          CCGACTAAGGCATTCTTAGCTGTTCGAGCGAGTTCAATATCCGTAAA-----CTTAAAGA 383
PO_B94          CCGACTAAGGCATTCTTAGCTGTTCGAGCGAGTTCAATATCCGTAAA-----CTTAAAGA 383
PO_B30          CCGACTAAGGCATTCTTAGCTGTTCGAGCGAGTTCAATATCCGTAAA-----CTTAAAGA 387
                .* : : ..  *:: * . * .:  . *  : : * : :**.   .*     .: ..: :

PO_B76          GCTCGCTTCAATGCTTGAGATGTATGTTTCAGGCGTTTCGACACGTAAAGTTTCAAAGAT 467
PO_B84          GCTCGCTTCAATGCTTGAGATGTATGTTTCAGGCGTTTCGACACGTAAAGTTTCAAAGAT 467
PO_B74          GCTCGCTTCAATGCTTGAGATGTATGTTTCAGGCGTTTCGACACGTAAAGTTTCAAAGAT 468
PO_B32          ACAGCCTTAAAATCAT-CACAGATTTCTTTTCGATCTTTGACACGTACTTTATGAGCAAT 463
PO_B80          ACAGCCTTAAAATCAT-CACAGATTTCTTTTCGATCTTTGACACGTACTTTATGAGCAAT 463
PO_B70          TCG-CTTTTACTGCTT-CTCTAAACGGTTTTGAATTCTTTTTTGGAATGTAACTGAAGAT 441
PO_B94          TCG-CTTTTACTGCTT-CTCTAAACGGTTTTGAATTCTTTTTTGGAATGTAACTGAAGAT 441
PO_B30          TCG-CTTTTACTGCTT-CTCTAAACGGTTTTGAATTCTTTTTTGGAATGTAACTGAAGAT 445
                 *    ** *.: *:* . .:.::   ** : ..   *  : : *:*   ::  ....**
```

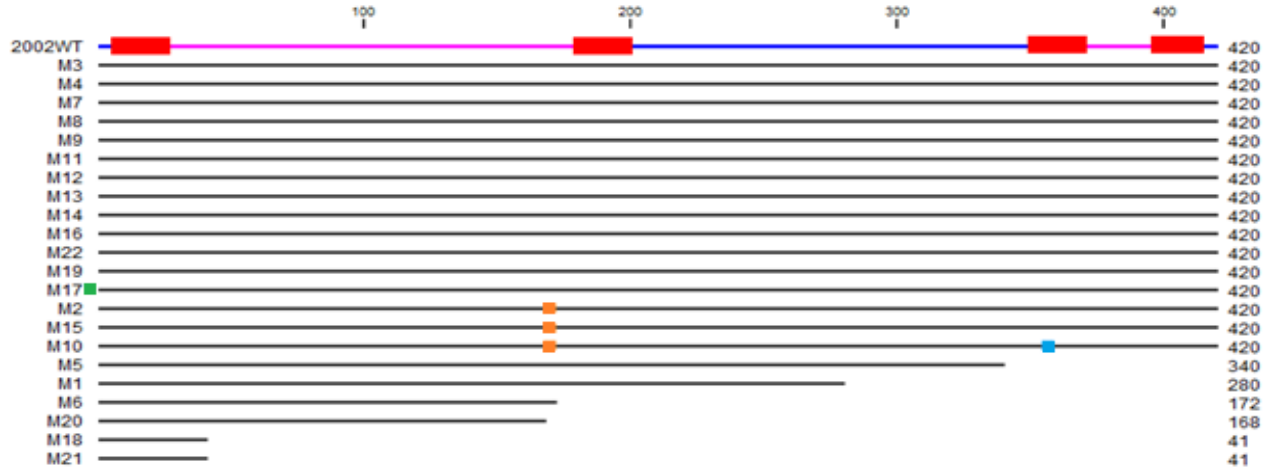
```
PO_B76          TGTTGAAGAGCTATG----TGGAAAATCGGTATCGAAATCTTTTG--TTTCTAGCCTGAC 521
PO_B84          TGTTGAAGAGCTATG----TGGAAAATCGGTATCGAAATCTTTTG--TTTCTAGCCTGAC 521
PO_B74          TGTTGAAGAGCTATG----TGGAAAATCGGTATCGAAATCTTTTG--TTTCTAGCCTGAC 522
PO_B32          ATTACGGGAAATGTGAACACAGCAATGTTGGTATTTA-GCTTTAG---GATAAATCTGGT 519
PO_B80          ATTACGGGAAATGTGAACACAGCAATGTTGGTATTTA-GCTTTAG---GATAAATCTGGT 519
PO_B70          GTTTCTTAAAAAATGGACCTGGCATCTCTGCCAACTT-GCGTTGG--TAAATG-ACTTAC 497
PO_B94          GTTTCTTAAAAAATGGACCTGGCATCTCTGCCAACTT-GCGTTGG--TAAATG-ACTTAC 497
PO_B30          GTTTCTTAAAAAATGGACCTGGCATCTCTGCCAACTTTGCGTTGGGTAAAATGAACTTAC 505
                  *:   .*..:.**     .*.*:    *  :  ::  * ** *    : :.  ** .

PO_B76          TGAGCAG-TTAGACCCGATG-----GTCAACGAATGGCAG-----AACCG-TTCACTCTC 569
PO_B84          TGAGCAG-TTAGACCCGATG-----GTCAACGAATGGCAG-----AACCG-TTCACTCTC 569
PO_B74          TGAGCAG-TTAGACCCGATG-----GTCAACGAATGGCAG-----AACCG-TTCACTCTC 570
PO_B32          GAATCG---TATCTTTCATG----CCTTTTAAGCCATCGGTAATAAACAGTAAGACATCA 572
PO_B80          GAATCG---TATCTTTCATG----CCTTTTAAGCCATCGGTAATAAACAGTAAGACATCA 572
PO_B70          GAATTG---CAGACACTAGG---CCTTTATGGG-CATCAG-----AAATG-ATGAG-TTC 543
PO_B94          GAATTG---CAGACACTAGG---CCTTTATGGG-CATCAG-----AAATG-ATGAG-TTC 543
PO_B30          GAAATTTGCAAGACACCTAGGGCCTTTTATGGGGCATCAAGA---AAATGAATGAGTTTC 562
                 .*       * .    : *      * :: ..  . *..     **. * :: *    .

PO_B76          AGGTACGAGTTATCCTTATCTGATGACTGATGTTCTCTACATAAAAGTCCGTGAGGACCA 629
PO_B84          AGGTACGAGTTATCCTTATCTGATGACTGATGTTCTCTACATAAAAGTCCGTGAGGACCA 629
PO_B74          AGGTACGAGTTATCCTTATCTGATGACTGATGTTCTCTACATAAAAGTCCGTGAGGACCA 630
PO_B32          TG-AACCCCTCTGGAGTAAATG-TCCTGTAACAACTCATTCCAAACGTAGGTCGATTCAG 630
PO_B80          TG-AACCCCTCTGGAGTAAATG-TCCTGTAACAACTCATTCCAAACGTAGGTCGATTCAG 630
PO_B70          TG--TCCCCTGTAG--ACCGCG-TTCTTTT--AAGTATT-CAAAG--AAGATGG--ACCA 591
PO_B94          TG--TCCCCTGTAG--ACCGCG-TTCTTTT--AAGTATT-CAAAG--AAGATGG--ACCA 591
PO_B30          TTGTCCCCCTGTAAGACCCGCGTTTCTTTTTAAAGTATTTCAAAAGAAAGATGGGACCCA 622
                :    * . * :     .   * * .   :  :: *.:: . **.  :. .* .   *..

PO_B76          TCGAGTGCTTTCTAAAG-CTGCCATATTG---CGATCGGGATAACAGAAGG---------- 676
PO_B84          TCGAGTGCTTTCTAAAAGCTGCCATATTG---CGATCGGGATAACAGAAGG---------- 677
PO_B74          TCGAGTGCTTTCTAAAG-CTGCCATATTG---CGATCGGGATAACAGAAGG---------- 677
PO_B32          TCG-GAGCAATCGCATAA-CTTAGTACT-------TCTTT-------AGTG---------- 665
PO_B80          TCG-GAGCAATCGCATAA-CTTAGTACT-------TCTTT-------AGTG---------- 665
PO_B70          TG---TGTCATCACTTT--CTTCATTTT----GGATCATG--AAGCCAATG-----ATTT 635
PO_B94          TG---TGTCATCACTTT--CTTCATTTT----GGATCATG--AAGCCAATG-----ATTT 635
PO_B30          TGGTGTTCCATCACTTTTCTTTCATTTTTTTGGAATCATTGAAAGCCCAAATGGATTTTC 682
                *    :   :** .::      ..*: *       **          .. .

PO_B76          ---------TGGTGATCGTG-AAATCATTGGCTTCATG--ATCCAAAATGAAGAAAGTGA 724
PO_B84          ---------TGGTGATCGTG-AAATCATTGGCTTCATG--ATCCAAAATGAAGAAAGTGA 725
PO_B74          ---------TGGTGATCGGG-AAACCAT-GGCTTCCTG--ATCCAAAATGAAGAAAGGGA 724
PO_B32          ---------CCGTCTTCTCGTATACCAATGGCAATATAGATTGCTTCTTTGGAGACGG-T 715
PO_B80          ---------CCGTCTTCTCGTATACCAATGGCAATATAGATTGCTTCTTTGGAGACGG-T 715
PO_B70          CACG-ATCACCACCTTCTGTTATCCCGATCGCAATATGGCAGCTTTTAGAAA-CACTCGA 693
PO_B94          CACG-ATCACCACCTTCTGTTATCCCGATCGCAATATGGCAGCTTTTAGAAAGCACTCGA 694
PO_B30          CACGAATCCACCCACCCTTTCTTGTTTATCCCCGAAT--CCGCAAATAATGGGCAAGCTT 740
                           *    ::    :   *   .*       :: :  ..  *.   :

PO_B76          TGACACATGGTCCATCTTCTTTGAAT------ACTTAAAAGAACGCGGTCTACAGGGGAC 778
PO_B84          TGACACATGGTCCATCTTCTTTGAAT------ACTTAAAAGAACGCGGTCTACAGGGGAC 779
PO_B74          TGCCCCAGG-TCCATCTTCTTTGAAT------ACTTAAAAGAACGCGGTTTACAGGGGAC 777
PO_B32          TTGCCGTTTTAAAGGAATATAAGTAG-----CATCCATGAAAATAGCGACATACTTATCA 770
PO_B80          TTGCCGTTTTAAAGGAATATAAGTAG-----CATCCATGAAAATAGCGACATACTTATCA 770
PO_B70          TGG-CCTCACGGACTTT-ATGTAAAG-ACATCAGTCA-CAGATAAGGA-AACTCGTAC-T 747
PO_B94          TGGTCCTCACGGACTTTTATGTAGAGAACATCAGTCATCAGATAAGGATAACTCGTACCT 754
PO_B30          TTTTAAAAAAAGCCACCTCCGAATGG---GTCCCTCCACGGGAACTTT--TTTAATGGTA 795
                *   . :     .     .  :. .       .   .  ...:      :  .  .

PO_B76          AGAACTCATCATTTCTGATGCCCATAAAGGCCTAGTGTCTGCAATTCGTAAGTCATTT-C 837
PO_B84          AGAACTCATCATTTCTGATGCCCATAAAGGCCTAGTGTCTGCAATTCGTAAGTCATTTAC 839
PO_B74          AGAATTCATCATTTCTGATGCCCATAAAGGCCTAGTGTCTGCAATTCGTAAGTCATTTAC 837
PO_B32          TTTAGAGTTCTAACTTTAAATGCATTTACTTCTTCAGTCAGCACTTTAGTCATGTTGG-A 829
PO_B80          TTTAGAGTTCTAACTTTAAATGCATTTACTTCTTCAGTCAGCACTTTAGTCATGTTGG-A 829
PO_B70          GAAAG-GAACGGTTCTGCCTTTC-TTGACCATCGG-TCTAACTGTTCAGTCAGGCTAG-A 803
PO_B94          GAGAGTGAACGGTTCTGCCATTCGTTGACCATCGGGTCTAACTGTTCAGTCAGGCTAG-A 813
PO_B30          AGAAGGAAACCTTCCAG--TTCCATTGACCATCGGGTCTAACTGCTCAGTCAGGCTAG-A 852
                   *   ::*  :  :      * *: *           :.*:  * . :..   *   .

PO_B76          CAACGCAAGTTGGCAGAGATGCCAGGTCCATTTTTTAAGAAACATCTTCAGTTACATTCC 897
PO_B84          CACCGCAAGTTGCCAGAGATGCCAGGTCCATTTTTTAAGAAACATCTTCAGTTACATTCC 899
PO_B74          CAACGCAAGTTGGCAGAGATGCCAGGTCCATTTTTTAAGAAACATCTTCAGTTACATTCC 897
PO_B32          CATGGTTTGGGGGCGTGTAGTGATGCCCATACATTTTCTCAATCAAGTCTGCGATTTCCG 889
PO_B80          CATGGTTTGGGG-CGTGTAGTGATGCCCATACATTTTCCCAATCAAGTCTGCGATTTCCG 888
PO_B70          AACA-AAAGATTTCGATACCGATTTTCCACATAGTTTTTCAACAATCTTTGAAACTTTAC 862
PO_B94          AACA-AAAGATTTCGATACCGATTTTCCACATAGCTCTTCAACAATCTTTGAAACTTTAC 872
PO_B30          AACA-AAAGATTTCGATACCGATTTTCCACATAGCTCTTCAACAATCTTTGAAACTTTAC 911
                .*    ::*    *.   .    :   *. : :  *   .** .:  * :*  * :* .

PO_B76          AAAAAAGAAT-TCAAAACCGTTTAGAGAAGCAGTAAAAGCGATCTTTAAGTTTACG---- 952
PO_B84          AAAAAAGAAT-TCAAAACCGTTTAGAGAAGCAGTAAAAGCGATCTTTAAGTTTACG---- 954
PO_B74          AAAAAAGAAT-TCAAAACCGTTTAGAGAAGCAGTAAAAGCGATCTTTAAGTTTACG---- 952
PO_B32          ACATGGTAATACCTTTTCGGAAGAGGTGAATGACGGTCTC---CTCTAAGGTATCATTCG 946
PO_B80          ACATGGTAATACCTTTTCGGAAAAAGTGAATGACGGTCTC---CTCTAAGGTATCATTCG 945
```

```
PO_B70          GTGTCGAAACGCCTGAAACATACATCTCAAGCATTGAAGCGAGCAGTGCCTTTTCATTTC 922
PO_B94          GTGTCGAAACGCCTGAAACATACATCTCAAGCATTGAAGCGAGCAGTGCCTTTTCATTTC 932
PO_B30          GTGTCGAAACGCCTGAAACATACATCTCAAGCATTGAAGCGAGCAGTGCCTTTTCATTTC 971
                . .: . **   *: ::. .:: *    *.  .  .:. *   *: *..  *::*.

PO_B76          ---GATATTGAACTCGCTCGAACAGCT----AAGAATGCCTTAGTCGGTGAATATATCGA 1005
PO_B84          ---GATATTGAACTCGCTCGAACAGCT----AAGAATGCCTTAGTCGGTGAATATATCGA 1007
PO_B74          ---GATATTGAACTCGCTCGAACAGCT----AAGAATGCCTTAGTCGGTGAATATATCGA 1005
PO_B32          TCCGTCTATAAGCAGGAACAGTCTGTTGCTTGAACTCGCCGTTGCGGTCGCGCGGAATCT 1006
PO_B80          TCCGTCTATAAGCAGGAACAGTCTGTTGCTTGAACTCGCCGTTGCGGTCGCGCGGAATCT 1005
PO_B70          GCTGATAACGCTCAAACACCGTCGGTG---AAAATTCACCATCACGTGTTCTAGGCACTT 979
PO_B94          GCTGATAACGCTCAAACACCGTCGGTG---AAAATTCACCATCACGTGTTCTAGGCACTT 989
PO_B30          GCTGATAACGCTCAAACACCGTCGGTG---AAAATTCACCATCACGTGTTCTAGGCACTT 1028
                   *: :: .. *: ..:* .:* *      .*. : .** * .      .  . .:  :

PO_B76          CCAGTCTAAATATACAAAAGCTTG--CGAAACATTGG----ATAATGGCTTCGAAGA--- 1056
PO_B84          CCAGTCTAAATATACAAAAGCTTG--CGAAACATTGG----ATAATGGCTTCGAAGA--- 1058
PO_B74          CCAGTCTAAATATACAAAAGCTTG--CGAAACATTGG----ATAATGGCTTCGAAGA--- 1056
PO_B32          GGAGATGAAGT-TCCCCGTACTCGGTCTTGACCGTACGGTCATAGGAGCCGTTACGAGAA 1065
PO_B80          GGAGATGAAGT-TCCCCGTACTCGGTCTTGACCGTACGGTCATAGGAGCCGTTACGAGAA 1064
PO_B70          TTAATTCGAGTGTGCCCACACGAG-TCGTAAAGTCTCGCTCATAATAGCCATTTCTT--- 1035
PO_B94          TTAATTCGAGTGTGCCCACACGAG-TCGTAAAGTCTCGCTCATAATAGCCATTTCTT--- 1045
PO_B30          TTAATTCGAGTGTGCCCACACGAG-TCGTAAAGTCTCGCTCATAATAGCCATTTCTT--- 1084
                  *.:  .*.* * *... .*  *  * :.*.          ***. .**    :. :

PO_B76          TGCCTTTCAATACACGGTTATCGGAAATGGTCACAATCG-GCTAAAAAGCACCAACCTTC 1115
PO_B84          TGCCTTTCAATACACGGTTATCGGAAATGGTCACAATCG-GCTAAAAAGCACCAACCTTC 1117
PO_B74          TGCCTTTCAATACACGGTTATCGGAAATGGTCACAATCG-GCTAAAAAGCACCAACCTTC 1115
PO_B32          TTACCCGTGTTAAAACCAATACGATCGTACTTTTCGTAATCGAGGAAAGCCGTTAACTCC 1125
PO_B80          TTACCCGTGTTAAAACCAATACGATCGTACTTTTCGTAATCGAGGAAAGCCGTTAACTCC 1124
PO_B70          TGACTCTGACGACTTTCTGAACGTTCATAGTCATCGGCTTGAATATATTCTGTTCGTTGA 1095
PO_B94          TGACTCTGACGACTTTCTGAACGTTCATAGTCATCGGCTTGAATATATTCTGTTCGTTGA 1105
PO_B30          TGACTCTGACGACTTTCTGAACGTTCATAGTCATCGGCTTGAATATATTCTGTTCGTTGA 1144
                * .*    .  *.:   : ::** :..*. * : .. .    : .:*: *   :.  * .

PO_B76          TTGAACGACTGAACCAGGAAGTCCGCAG-----AAGAGAAAAGATTATTCGGATTTTTCC 1170
PO_B84          TTGAACGACTGAACCAGGAAGTCCGCAG-----AAGAGAAAAGATTATTCGGATTTTTCC 1172
PO_B74          TTGAACGACTGAACCAGGAAGTCCGCAG-----AAGAGAAAAGATTATTCGGATTTTTCC 1170
PO_B32          GTTTTGAGAAGCGTATTGACAGCTTT-------TTCCAAATGGACACGGAATAATTCATT 1178
PO_B80          GTTTTGAGAAGCGTATTGACAGCTTT-------TTCCAAATGGACACGGAATAATTCATT 1177
PO_B70          TTTTCCATCAATTGGTTGAAAACAGTGGTTAAAATATTTTTAGAAACATCATCTTTTACC 1155
PO_B94          TTTTCCATCAATTGGTTGAAAACAGTGGTTAAAATATTTTTAGAAACATCATCTTTTACC 1165
PO_B30          TTTTCCATCAATTGGTTGAAAACAGTGGTTAAAATATTTTTAGAAACATCATCTTTTACC 1204
                 * :  . .:.    : **.. *          ::   :::.** :.  .. .:** :

PO_B76          CAACCGAACGTCTGCCAATCGATTAATTGGAG---CTGTCCTTATGG----ACCTTCATG 1223
PO_B84          CAACCGAACGTCTGCCAATCGATTAATTGGAG---CTGTCCTTATGG----ACCTTCATG 1225
PO_B74          CAACCGAACGTCTGCCAATCGATTAATTGGAG---CTGTCCTTATGG----ACCTTCATG 1223
PO_B32          CAAATCGCCTTTGTTCGCTAGAGTCTTTAGAATTTCTGTAGTAAAA-----TCGTTCATA 1233
PO_B80          CAAATCGCCTTTGTTCGCTAGAGTCTTTAGAATTTCTGTAGTAAAA-----TCGTTCATA 1232
PO_B70          GAATGTTCAATAATACTTTGAACCTCTTCGTTGTTCAGTGTAAAATGTACTTGGGTCATG 1215
PO_B94          GAATGTTCAATAATACTTTGAACCTCTTCGTTGTTCAGTGTAAAATGTACTTGGGTCATG 1225
PO_B30          GAATGTTCAATAATACTTTGAACCTCTTCGTTGTTCAGTGTAAAATGTACTTGGGTCATG 1264
                 **    .. *    *  * .*    ** *:    *:**  ::*:      :   ****.

PO_B76          ACGAATGGCTCAGTTCTACAAGAAAATAT--ATTAAGTTTGATCAATGAGAGACCGGTAA 1281
PO_B84          ACGAATGGCTCAGTTCTACAAGAAAATAT--ATTAAGTTTGATCAATGAGAGACCGGTAA 1283
PO_B74          ACGAATGGCTCAGTTCTACAAGAAAATAT--ATTAAGTTTGATCAATGAGAGACCGGTAA 1281
PO_B32          TGGAAGTCCTCTTTTCTGTGAATGTGTTG-TGGTAACTCTATTCTACAGAAGGG-----A 1287
PO_B80          TGGAAGTCCTCTTTTCTGTGAATGTGTTG-TGGTAACTCTATTCTACAGAAGGG-----A 1286
PO_B70          T-AACGTCCTCCTGGGTATGTTTTTGTGGTTAAAAACATTGTACCGTAAAAGGGCTGTTA 1274
PO_B94          T-AACGTCCTCCTGGGTATGTTTTTGTGGTTAAAAACATTGTACCGTAAAAGGGCTGTTA 1284
PO_B30          T-AACGTCCTCCTGGGTATGTTTTTGTGGTTAAAAACATTGTACCGTAAAAGGGCTGTTA 1323
                : .*.   ***     *. .: : :.*    . :** : *.::* . ...**.      *

PO_B76          AACATTGTATAGTAT-TTTACACAGGATTATGGACTTGACT------------ 1321
PO_B84          AACATTGTATAGTAT-TTTACACAGGATTATGGACTTGACT------------ 1323
PO_B74          AACATTGTATAGTAT-TTTACACAGGATTATGGACTTGACT------------ 1321
PO_B32          CTT--CCTTTTTTCTATTTACACAAAATATTTTACACTCTC------------ 1326
PO_B80          CTT--CCTTTTTTCTATTTACACAAAATATTTTACACTCTCG----------- 1326
PO_B70          CATGGCCTTTTATCT-TTTACACAATTATATGGACTTTATC------------ 1314
PO_B94          CATGGCCTTTTATCT-TTTACACAATTATATGGACTTTATCAAAACGATACTA 1336
PO_B30          CATGGCCTTTTATCT-TTTACACAATTATATGGACTTTATCCATT-------- 1367
                .:     *:*: *.* ********. ::::*  **:  .
```

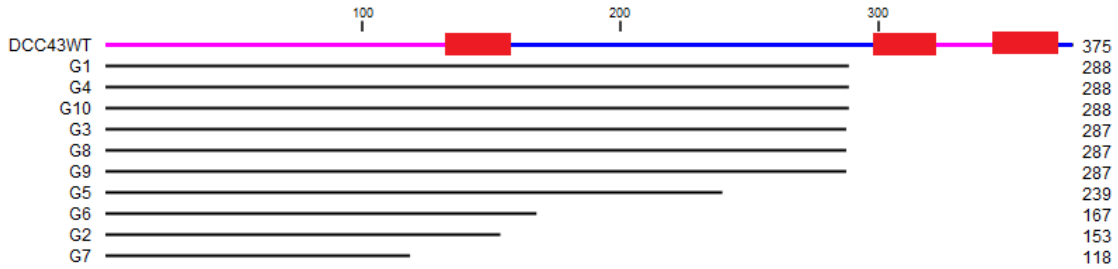
Visualization of truncated RseP proteins



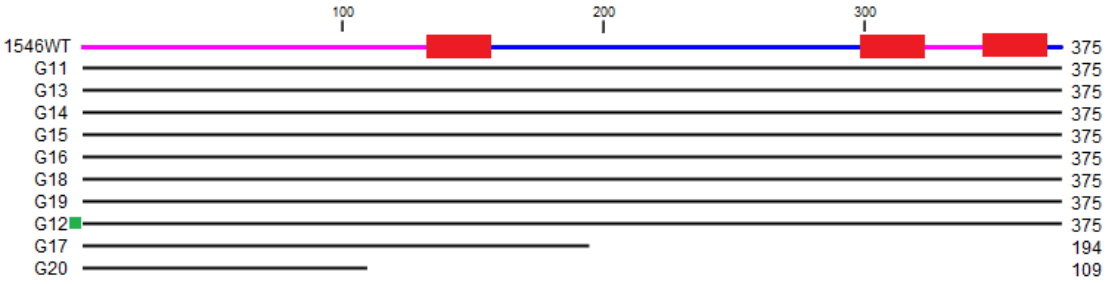
A10: The RseP protein lengths of all the sequenced *E. faecalis* 3358 mutants as displayed with CLC workbench. The wild type is displayed at the top with transmembrane regions (red), regions inside the membrane (blue) and outside the membrane (pink) visualized as predicted by TMHMM server v.2.0 (http://www.cbs.dtu.dk/services/TMHMM). These are only predictions based on the best result according to the algorithm, but interesting nevertheless.



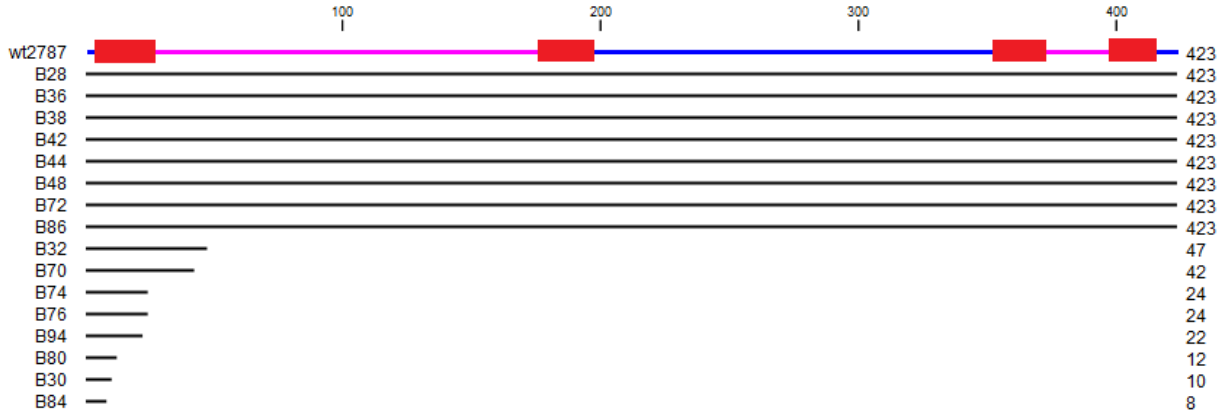
A11: The RseP protein lengths of the *P. acidilactici* 2002 mutants. The wild type is displayed with transmembrane regions (red), regions inside the membrane (blue) and outside the membrane (pink) as predicted by TMHMM. Green square: 9 bp upstream gene G>A mutation. Orange square: Arg171Lys. Blue square: Asn357Lys.



A12: The RseP protein lengths of the *L. garvieae* DCC43 mutants. The wild type is displayed with transmembrane regions (red), regions inside the membrane (blue) and outside the membrane (pink) as predicted by TMHMM.



A13: The RseP protein lengths of the *L. garvieae* 1546 mutants. The wild type is displayed with transmembrane regions (red), regions inside the membrane (blue) and outside the membrane (pink) as predicted by TMHMM. Green square: marks the location of a 2 bp upstream gene mutation.



A14: The RseP protein lengths of the *E. faecium* 2787 mutants. The wild type is displayed with transmembrane regions (red), regions inside the membrane (blue) and outside the membrane (pink) as predicted by TMHMM. Due to the transposons at the beginning of the gene sequence (making the PCR amplicons twice as large), the proteins are severely impacted and terminated at an early stage.

SNP detection: mutational consequences on EcsAB

A15: MSA of EcsA 3358

```
3358wt_EcsA      MSLTIEHLTGGYGHIPVLKDINFDVKSGEMVGLIGLNGAGKSTTIKNIIGLLTPQKGKIM 60
E3_EcsA          MSLTIEHLTGGYGHIPVLKDINFDVKSGEMVGLIGLNGAGKSTTIKNIIGLLTPQKGKIM 60
E25_EcsA         MSLTIEHLTGGYGHIPVLKDINFDVKSGEMVGLIGLNGAGKSTTIKNIIGLLTPQKGKIM 60
E4_EcsA          MSLTIEHLTGGYGHIPVLKDINFDVKSGEMVGLIGLNGAGKSTTIKNIIGLLTPQKGKIM 60
E24_EcsA         MSLTIEHLTGGYGHIPVLKDINFDVKSGEMVGLIGLNGAGKSTTIKNIIGLLTPQKGKIM 60
E2_EcsA          MSLTIEHLTGGYGHIPVLKDINFDVKSGEMVGLIGLNGAGKSTTIKNIIGLLTPQKGKIM 60
E8_EcsA          MSLTIEHLTGGYGHIPVLKDINFDVKSGEMVGLIGLNGAGKSTTIKNIIGLLTPQKGKIM 60
E11_EcsA         MSLTIEHLTGGYGHIPVLKDINFDVKSGEMVGLIGLNGAGKSTTIKNIIGLLTPQKGKIM 60
E12_EcsA         MSLTIEHLTGGYGHIPVLKDINFDVKSGEMVGLIGLNGAGKSTTIKNIIGLLTPQKGKIM 60
E14_EcsA         MSLTIEHLTGGYGHIPVLKDINFDVKSGEMVGLIGLNGAGKSTTIKNIIGLLTPQKGKIM 60
E26_EcsA         MSLTIEHLTGGYGHIPVLKDINFDVKSGEMVGLIGLNGAGKSTTIKNIIGLLTPQKGKIM 60
                 ************************************************************

3358wt_EcsA      IDGETLQQAPEEYRKKIGYIPETPSLYEELTLKEHIEVTALAYDIPLEEAFKRAEPLLKT 120
E3_EcsA          IDGETLQQAPEEYRKKIGYIPETPSLYEELTLKEHIEVTALAYDIPLEEAFKRAEPLLKT 120
E25_EcsA         IDGETLQQAPEEYRKKIGYIPETPSLYEELTLKEHIEVTALAYDIPLEEAFKRAEPLLKT 120
E4_EcsA          IDGETLQQAPEEYRKKIGYIPETPSLYEELTLKEHIEVTALAYDIPLEEAFKRAEPLLKT 120
E24_EcsA         IDGETLQQAPEEYRKKIGYIPETPSLYEELTLKEHIEVTALAYDIPLEEAFKRAEPLLKT 120
E2_EcsA          IDGETLQQAPEEYRKKIGYIPETPSLYEELTLKEHIEVTALAYDIPLEEAFKRAEPLLKT 120
E8_EcsA          IDGETLQQAPEEYRKKIGYIPETPSLYEELTLKEHIEVTALAYDIPLEEAFKRAEPLLKT 120
E11_EcsA         IDGETLQQAPEEYRKKIGYIPETPSLYEELTLKEHIEVTALAYDIPLEEAFKRAEPLLKT 120
E12_EcsA         IDGETLQQAPEEYRKKIGYIPETPSLYEELTLKEHIEVTALAYDIPLEEAFKRAEPLLKT 120
E14_EcsA         IDGETLQQAPEEYRKKIGYIPETPSLYEELTLKEHIEVTALAYDIPLEEAFKRAEPLLKT 120
E26_EcsA         IDGETLQQAPEEYRKKIGYIPETPSLYEELTLKEHIEVTALAYDIPLEEAFKRAEPLLKT 120
                 ************************************************************

3358wt_EcsA      FRLDNKLEWFPANFSKGMKQKVMVLCAFLIEPSLYIIDEPFLGLDPLAIHALLELMDTMR 180
E3_EcsA          FRLDNKLEWFPANFSKGMKQKVMVLCAFLIEPSLYIIDEPFLGLDPLAIHALLELMDTMR 180
E25_EcsA         FRLDNKLEWFPANFSKGMKQKVMVLCAFLIEPSLYIIDEPFLGLDPLAIHALLELMDTMR 180
E4_EcsA          FRLDNKLEWFPANFSKGMKQKVMVLCAFLIEPSLYIIDEPFLGLDPLAIHALLELMDTMR 180
E24_EcsA         FRLDNKLEWFPANFSKGMKQKVMVLCAFLIEPSLYIIDEPFLGLDPLAIHALLELMDTMR 180
E2_EcsA          FRLDNKL------------------------------------------------------- 127
E8_EcsA          FRLDNKL------------------------------------------------------- 127
E11_EcsA         FRLDNKL------------------------------------------------------- 127
E12_EcsA         FRLDNKL------------------------------------------------------- 127
E14_EcsA         FRLDNKL------------------------------------------------------- 127
E26_EcsA         FRLDNKL------------------------------------------------------- 127
                 *******

3358wt_EcsA      KQGAAILMSTHILATAEKYCDRFVVLHEGKLRANGTMAELRAEFNLPESSLDDIYLALTK 240
E3_EcsA          KQGAAILMSTHILATAEKYCDRFVVLHEGKLRANGTMAELRAEFNLPESSLDDIYLALTK 240
E25_EcsA         KQGAAILMSTHILATAEKYCDRFVVLHEGKLRANGTMAELRAEFNLPESSLDDIYLALTK 240
E4_EcsA          KQGAAILMSTHILATAEKYCDRFVVLHEGKLRANGTMAELRAEFNLPESSLDDIYLALTK 240
E24_EcsA         KQGAAILMSTHILATAEKYCDRFVVLHEGKLRANGTMAELRAEFNLPESSLDDIYLALTK 240
E2_EcsA          ------------------------------------------------------------
E8_EcsA          ------------------------------------------------------------
E11_EcsA         ------------------------------------------------------------
E12_EcsA         ------------------------------------------------------------
E14_EcsA         ------------------------------------------------------------
E26_EcsA         ------------------------------------------------------------


3358wt_EcsA      EEKVGX 246
E3_EcsA          EEKVGX 246
E25_EcsA         EEKVGX 246
E4_EcsA          EEKVGX 246
E24_EcsA         EEKVGX 246
E2_EcsA          ------
E8_EcsA          ------
E11_EcsA         ------
E12_EcsA         ------
E14_EcsA         ------
E26_EcsA         ------
```

A16: MSA of EcsA 2002

```
2002wt_EcsA      LTLKIEHLTGGYTQTPVIKDINLEVHEGELVGLIGLNGAGKSTTINHVIGLLQPMQGTIS 60
M1_EcsA          LTLKIEHLTGGYTQTPVIKDINLEVHEGELVGLIGLNGAGKSTTINHVIGLLQPMQGTIS 60
M5_EcsA          LTLKIEHLTGGYTQTPVIKDINLEVHEGELVGLIGLNGAGKSTTINHVIGLLQPMQGTIS 60
M4_EcsA          LTLKIEHLTGGYTQTPVIKDINLEVHEGELVGLIGLNGAGKSTTINHVIGLLQPMQGTIS 60
M7_EcsA          LTLKIEHLTGGYTQTPVIKDINLEVHEGELVGLIGLNGAGKSTTINHVIGLLQPMQGTIS 60
M8_EcsA          LTLKIEHLTGGYTQTPVIKDINLEVHEGELVGLIGLNGAGKSTTINHVIGLLQPMQGTIS 60
M11_EcsA         LTLKIEHLTGGYTQTPVIKDINLEVHEGELVGLIGLNGAGKSTTINHVIGLLQPMQGTIS 60
M12_EcsA         LTLKIEHLTGGYTQTPVIKDINLEVHEGELVGLIGLNGAGKSTTINHVIGLLQPMQGTIS 60
M16_EcsA         LTLKIEHLTGGYTQTPVIKDINLEVHEGELVGLIGLNGAGKSTTINHVIGLLQPMQGTIS 60
M19_EcsA         LTLKIEHLTGGYTQTPVIKDINLEVHEGELVGLIGLNGAGKSTTINHVIGLLQPMQGTIS 60
M22_EcsA         LTLKIEHLTGGYTQTPVIKDINLEVHEGELVGLIGLNGAGKSTTINHVIGLLQPMQGTIS 60
                 ************************************************************

2002wt_EcsA      LNNIQLSQQPAAYKAQLAYIPEMPVLYDELTLKEHLELTIKAYDLDEKVAWKNAMELLEK 120
M1_EcsA          LNNIQLSQQPAAYKAQLAYIPEMPVLYDELTLKEHLELTIKAYDLDEKVAWKNAMELLEK 120
M5_EcsA          LNNIQLSQQPAAYKAQLAYIPEMPVLYDELTLKEHLELTIKAYDLDEKVAWKNAMELLEK 120
M4_EcsA          LNNIQLSQQPAAYKAQLAYIPEMPVLYDELTLKEHLELTIKAYDLDEKVAWKNAMELLEK 120
M7_EcsA          LNNIQLSQQPAAYKAQLAYIPEMPVLYDELTLKEHLELTIKAYDLDEKVAWKNAMELLEK 120
M8_EcsA          LNNIQLSQQPAAYKAQLAYIPEMPVLYDELTLKEHLELTIKAYDLDEKVAWKNAMELLEK 120
M11_EcsA         LNNIQLSQQPAAYKAQLAYIPEMPVLYDELTLKEHLELTIKAYDLDEKVAWKNAMELLEK 120
M12_EcsA         LNNIQLSQQPAAYKAQLAYIPEMPVLYDELTLKEHLELTIKAYDLDEKVAWKNAMELLEK 120
M16_EcsA         LNNIQLSQQPAAYKAQLAYIPEMPVLYDELTLKEHLELTIKAYDLDEKVAWKNAMELLEK 120
M19_EcsA         LNNIQLSQQPAAYKAQLAYIPEMPVLYDELTLKEHLELTIKAYDLDEKVAWKNAMELLEK 120
M22_EcsA         LNNIQLSQQPAAYKAQLAYIPEMPVLYDELTLKEHLELTIKAYDLDEKVAWKNAMELLEK 120
                 ************************************************************

2002wt_EcsA      FRLANKLEWFPANFSKGMRQKVMIVCAFMTDARLFVIDEPFLGLDPLAIDDLLGIINEKK 180
M1_EcsA          FRLANKLEWFPANFSKGMRQKVMIVCAFMTDARLFVIDEPFLGLDPLAIDDLLGIINEKK 180
M5_EcsA          FRLANKLEWFPANFSKGMRQKVMIVCAFMTDARLFVIDEPFLGLDPLAIDDLLGIINEKK 180
M4_EcsA          FRLANKLEWFPANFSKGMRQKVMIVCAFMTDARLFVIDEPFLGLDPLAIDDLLGIINEKK 180
M7_EcsA          FRLANKLEWFPANFSKGMRQKVMIVCAFMTDARLFVIDEPFLGLDPLAIDDLLGIINEKK 180
M8_EcsA          FRLANKLEWFPANFSKGMRQKVMIVCAFMTDARLFVIDEPFLGLDPLAIDDLLGIINEKK 180
M11_EcsA         FRLANKLEWFPANFSKGMRQKVMIVCAFMTDARLFVIDEPFLGLDPLAIDDLLGIINEKK 180
M12_EcsA         FRLANKLEWFPANFSKGMRQKVMIVCAFMTDARLFVIDEPFLGLDPLAIDDLLGIINEKK 180
M16_EcsA         FRLANKLEWFPANFSKGMRQKVMIVCAFMTDARLFVIDEPFLGLDPLAIDDLLGIINEKK 180
M19_EcsA         FRLANKLEWFPANFSKGMRQKVMIVCAFMTDARLFVIDEPFLGLDPLAIDDLLGIINEKK 180
M22_EcsA         FRLANKLEWFPANFSKGMRQKVMIVCAFMTDARLFVIDEPFLGLDPLAIDDLLGIINEKK 180
                 ************************************************************

2002wt_EcsA      KQGAAVLMSTHVLDTAEKYCDRFALLNEGQLKAVGTLGDFQTQFNAPTASLNDIYLSIAK 240
M1_EcsA          KQGAAVLMSTHVLDTAEKYCDRFALLNEGQLKAVGTLGDFQTQFNAPTASLNDIYLSIAK 240
M5_EcsA          KQGAAVLMSTHVLDTAEKYCDRFALLNEGQLKAVGTLGDFQTQFNAPTASLNDIYLSIAK 240
M4_EcsA          KQGAAVLMSTHVLDTAEKYCDRFALLNEGQLKAVGTLGDFQTQFNAPTASLNDIYLSIAK 240
M7_EcsA          KQGAAVLMSTHVLDTAEKYCDRFALLNEGQLKAVGTLGDFQTQFNAPTASLNDIYLSIAK 240
M8_EcsA          KQGAAVLMSTHVLDTAEKYCDRFALLNEGQLKAVGTLGDFQTQFNAPTASLNDIYLSIAK 240
M11_EcsA         KQGAAVLMSTHVLDTAEKYCDRFALLNEGQLKAVGTLGDFQTQFNAPTASLNDIYLSIAK 240
M12_EcsA         KQGAAVLMSTHVLDTAEKYCDRFALLNEGQLKAVGTLGDFQTQFNAPTASLNDIYLSIAK 240
M16_EcsA         KQGAAVLMSTHVLDTAEKYCDRFALLNEGQLKAVGTLGDFQTQFNAPTASLNDIYLSIAK 240
M19_EcsA         KQGAAVLMSTHVLDTAEKYCDRFALLNEGQLKAVGTLGDFQTQFNAPTASLNDIYLSIAK 240
M22_EcsA         NRGRRY---------------------------------------------------------- 186
                 ::*         *   ::    :  *.:*. : :.     : .:  *      *::.  :

2002wt_EcsA      GERHEX 246
M1_EcsA          GERHEX 246
M5_EcsA          GERHEX 246
M4_EcsA          GERHEX 246
M7_EcsA          GERHEX 246
M8_EcsA          GERHEX 246
M11_EcsA         GERHEX 246
M12_EcsA         GERHEX 246
M16_EcsA         GERHEX 246
M19_EcsA         GERHEX 246
M22_EcsA         ------
```

A17: MSA of EcsA 1546

```
1546wt_EcsA      VALKINNVTGGYFGNPVLEDVSFEIADGELVGLIGLNGAGKSTTIQEIMGLLTPYSGQIE 60
G12_EcsA         VALKINNVTGGYFGNPVLEDVSFEIADGELVGLIGLNGAGKSTTIQEIMGLLTPYSGQIE 60
G17_EcsA         VALKINNVTGGYFGNPVLEDVSFEIADGELVGLIGLNGAGKSTTIQEIMGLLTPYSGQIE 60
G20_EcsA         VALKINNVTGGYFGNPVLEDVSFEIADGELVGLIGLNGAGKSTTIQEIMGLLTPYSGQIE 60
G13_EcsA         VALKINNVTGGYFGNPVLEDVSFEIADGELVGLIGLNGAGKSTTIQEIMGLLTPYSGQIE 60
G18_EcsA         VALKINNVTGGYFGNPVLEDVSFEIADGELVGLIGLNGAGKSTTIQEIMGLLTPYSGQIE 60
G16_EcsA         VALKINNVTGGYFGNPVLEDVSFEIADGELVGLIGLNGAGKSTTIQEIMGLLTPYSGQIE 60
G19_EcsA         VALKINNVTGGYFGNPVLEDVSFEIADGELVGLIGLNGAGKSTTIQEIMGLLTPYSGQIE 60
G14_EcsA         VALKINNVTGGYFGNPVLEDVSFEIADGELVGLIGLNGAGKSTTIQEIMGLLTPYSGQIE 60
G11_EcsA         VALKINNVTGGYFGNPVLEDVSFEIADGELVGLIGLNGAGKSTTIQEIMGLLTPYSGQIE 60
G15_EcsA         VALKINNVTGGYFGNPVLEDVSFEIADGELVGLIGLNGAGKSTTIQEIMGLLTPYSG--- 57
                 ********************************************************

1546wt_EcsA      LDGITLQEDLEGYRKKIGFIPETPSLYEELTLREHIELTALAYGVPGEEAMPRAEELLKT 120
G12_EcsA         LDGITLQEDLEGYRKKIGFIPETPSLYEELTLREHIELTALAYGVPGEEAMPRAEELLKT 120
G17_EcsA         LDGITLQEDLEGYRKKIGFIPETPSLYEELTLREHIELTALAYGVPGEEAMPRAEELLKT 120
G20_EcsA         LDGITLQEDLEGYRKKIGFIPETPSLYEELTLREHIELTALAYGVPGEEAMPRAEELLKT 120
G13_EcsA         LDGITLQEDLEGYRKKIGFIPETPSLYEELTLREHIELTALAYGVPGEEAMPRAEELLKT 120
G18_EcsA         LDGITLQEDLEGYRKKIGFIPETPSLYEELTLREHIELTALAYGVPGEEAMPRAEELLKT 120
G16_EcsA         LDGITLQEDLEGYRKKIGFIPETPSLYEELTLREHIELTALAYGVPGEEAMPRAEELLKT 120
G19_EcsA         LDGITLQEDLEGYRKKIGFIPETPSLYEELTLREHIELTALAYGVPGEEAMPRAEELLKT 120
G14_EcsA         LDGITLQEDLEGYRKKIGFIPETPSLYEELTLREHIELTALAYGVPGEEAMPRAEELLKT 120
G11_EcsA         LD------------------------------------------------------------ 62
G15_EcsA         ------------------------------------------------------------

1546wt_EcsA      FRLEDKLDWFPVNFSKGMKQKVMIICAFLTDPSLYIIDEPFLGLDPLAIQDLITLMLSMR 180
G12_EcsA         FRLEDKLDWFPVNFSKGMKQKVMIICAFLTDPSLYIIDEPFLGLDPLAIQDLITLMLSMR 180
G17_EcsA         FRLEDKLDWFPVNFSKGMKQKVMIICAFLTDPSLYIIDEPFLGLDPLAIQDLITLMLSMR 180
G20_EcsA         FRLEDKLDWFPVNFSKGMKQKVMIICAFLTDPSLYIIDEPFLGLDPLAIQDLITLMLSMR 180
G13_EcsA         FRLEDKLDWFPVNFSKGMKQKVMIICAFLTDPSLYIIDEPFLGLDPLAIQDLITLMLSMR 180
G18_EcsA         FRLEDKLDWFPVNFSKGMKQKVMIICAFLTDPSLYIIDEPFLGLDPLAIQDLITLMLSMR 180
G16_EcsA         LRLEDKLDWFPVNFSKGMKQKVMIICAFLTDPSLYIIDEPFLGLDPLAIQDLITLMLSMR 180
G19_EcsA         FRLEDKLDWFPVNFSKGMKQKVMIICAFLTDPSLYIIDEPFLVLDPLAIQDLITLMLSMR 180
G14_EcsA         FRLEDKLDWFPVNFSKGMKQKVMIICAFLTDPSL--------------------------- 154
G11_EcsA         ------------------------------------------------------------
G15_EcsA         ------------------------------------------------------------

1546wt_EcsA      NAGASILMSTHILSTAEKFCDKFVILHEGKVLIAGTMEDLRAKFGKEDASLDEIYLELTK 240
G12_EcsA         NAGASILMSTHILSTAEKFCDKFVILHEGKVLIAGTMEDLRAKFGKEDASLDEIYLELTK 240
G17_EcsA         NAGASILMSTHILSTAEKFCDKFVILHEGKVLIAGTMEDLRAKFGKEDASLDEIYLELTK 240
G20_EcsA         NAGASILMSTHILSTAEKFCDKFVILHEGKVLIAGTMEDLRAKFGKEDASLDEIYLELTK 240
G13_EcsA         NAGASILMSTHILSTAEKFCDKFVILHEGKVLIAGTMEDLRAKFGKEDASLDEIYLELTK 240
G18_EcsA         NAGASILMSTHILSTAEKFCDKFVILHEGKVLIAGTMEDLRAKFGKEDASLDEIYLELTK 240
G16_EcsA         NAGASILMSTHILSTAEKFCDKFVILHEGKVLIAGTMEDLRAKFGKEDASLDEIYLELTK 240
G19_EcsA         NAGASILMSTHILSTAEKFCDKFVILHEGKVLIAGTMEDLRAKFGKEDASLDEIYLELTK 240
G14_EcsA         ------------------------------------------------------------
G11_EcsA         ------------------------------------------------------------
G15_EcsA         ------------------------------------------------------------

1546wt_EcsA      GVNAHEX 247
G12_EcsA         GVNAHEX 247
G17_EcsA         GVNAHEX 247
G20_EcsA         GVNAHEX 247
G13_EcsA         GVNAHEX 247
G18_EcsA         GVNAHEX 247
G16_EcsA         GVNAHEX 247
G19_EcsA         GVNAHEX 247
G14_EcsA         -------
G11_EcsA         -------
G15_EcsA         -------
```

A18: MSA of EcsB 2002

```
EcsB_2002wt      MSELWQQRAKTYQSEILKYLKYVLNDHLVLALLFFGGGLGLAYANWLKTLQPLPNYGLIL 60
M1_EcsB          MSELWQQRAKTYQSEILKYLKYVLNDHLVLALLFFGGGLGLAYANWLKTLQPLPNYGLIL 60
M5_EcsB          MSELWQQRAKTYQSEILKYLKYVLNDHLVLALLFFGGGLGLAYANWLKTLQPLPNYGLIL 60
TM22_EcsB        MSELWQQRAKTYQSEILKYLKYVLNDHLVLALLFFGGGLGLAYANWLKTLQPLPNYGLIL 60
M4_EcsB          MSELWQQRAKTYQSEILKYLKYVLNDHLVLALLFFGGGLGLAYANWLKTLQPLPNYGLIL 60
M16_EcsB         MSELWQQRAKTYQSEILKYLKYVLNDHLVLALLFFGGGLGLAYANWLKTLQPLPNYGLIL 60
M8_EcsB          MSELWQQRAKTYQSEILKYLKYVLNDHLVLALLFFGGGLGLAYANWLKTLQPLPNYGLIL 60
M12_EcsB         MSELWQQRAKTYQSEILKYLKYVLNDHLVLALLFFGGGLGLAYANWLKTLQPLPNYGLIL 60
M7_EcsB          MSELWQQRAKTYQSEILKYLKYVLNDHLVLALLFFGGGLG--------------------- 39
M19_EcsB         MSELWQQRAKTYQSEILKYLKYVLNDHLVLALLFFGAA---------------------- 37
M11_EcsB         MSELWQQRAKTYQSEILKYLKYVLNDHLVLALLFFGAA---------------------- 37
                 ***********************************

EcsB_2002wt      VATVSFLLTLPGQPVLLLKNADPVFLMSQEKNISAYLRKSLRRTLISVLIPVGIGSLLLF 120
M1_EcsB          VATVSFLLTLPGQPVLLLKNADPVFLMSQEKNISAYLRKSLRRTLISVLIPVGIGSLLLF 120
M5_EcsB          VATVSFLLTLPGQPVLLLKNADPVFLMSQEKNISAYLRKSLRRTLISVLIPVGIGSLLLF 120
TM22_EcsB        VATVSFLLTLPGQPVLLLKNADPVFLMSQEKNISAYLRKSLRRTLISVLIPVGIGSLLLF 120
M4_EcsB          VATVSFLLTLPGQPVLLLKNADPVFLMSQEKNISAYLRKSLRRTLISVLIPVGIGSLLLF 120
M16_EcsB         VATVSFLLTLPGQPVLLLKNADPVFLMSQEKNISAYLRKSLRRTLISVLIPVGIGSLLLF 120
M8_EcsB          VATVSFLLTLPGQPVLLLKNADPVFLMSQEKNISAYLRKSLRRTLISVLIPVGIGSLLLF 120
M12_EcsB         VATVSFLLTLPGQPVLLLKNADPVFLMSQEKNISAYLRKSLRRTLISVLIPVGIVACCFF 120
M19_EcsB         ------------------------------------------------------------
M11_EcsB         ------------------------------------------------------------
M7_EcsB          ------------------------------------------------------------

EcsB_2002wt      PLLALAYHQLFLVVGVILTVILGSFTNVLIRYQGYYFAGSRWDWKITLLIDAI 173
M1_EcsB          PLLALAYHQLFLVVGVILTVILGSFTNVLIRYQGYYFAGSRWDWKITLLIDAI 173
M5_EcsB          PLLALAYHQLFLVVGVILTVILGSFTNVLIRYQGYYFAGSRWDWKITLLIDAI 173
TM22_EcsB        PLLALAYHQLFLVVGVILTVILGSFTNVLIRYQGYYFAGSRWDWKITLLIDAI 173
M4_EcsB          PLLALAYHQLFLVVGVILTVILGSFTNVLIRYQGYYFAGSRWDWKITLLIDAI 173
M16_EcsB         PLLALAYHQLFLVVGVILTVILGSFTNVLIRYQGYYFAGSRWDWKITLLIDAI 173
M8_EcsB          PLLALAYHQLFLVVGVILTVILGSFTNVLIRYQGYYFAGSRWDWKITLLIDAI 173
M12_EcsB         HC-------------------------------------------------------
M19_EcsB         ---------------------------------------------------------
M11_EcsB         ---------------------------------------------------------
M7_EcsB          ---------------------------------------------------------

EcsB_2002wt      IVSLIIRFPLIAAVLMLGWLVLEWQRLNRLMEKGSFNWQAAVDAEQKRMGRLYRFFSLF 232
M1_EcsB          IVSLIIRFPLIAAVLMLGWLVLEWQRLNRLMEKGSFNWQAAVDAEQKRMGRLYRFFSLF 232
M5_EcsB          IVSLIIRFPLIAAVLMLGWLVLEWQRLNRLMEKGSFNWQAAVDAEQKRMGRLYRFFSLF 232
TM22_EcsB        IVSLIIRFPLIAAVLMLGWLVLEWQRLNRLMEKGSFNWQAAVDAEQKRMGRLYRFFSLF 232
M4_EcsB          IVSLIIRFPLIAAVLMLGWLVLEWQRLNRLMEKGSFNWQAAVDAEQKRMGRLYRFFSLF 232
M16_EcsB         IVSLIIRFPLIAAVLMLGWLVLEWQRLNRLMEKGSFNWQAAVDAEQKRMGRLYRFFSLF 232
M8_EcsB          IVSLIIRFPLIAAVLMLGWLVLEWQRLNRLMEKGSFNWQGGRCG-TKANGTVIPLLQLI 231
M12_EcsB         ------------------------------------------------------------
M19_EcsB         ------------------------------------------------------------
M11_EcsB         ------------------------------------------------------------
M7_EcsB          ------------------------------------------------------------

EcsB_2002wt      TTVKNLPVKAKRRRYADGLVKFLAGKSSMFSYLYSIMIVRSGDLGSLMIRLLIVGMLIVA 292
M1_EcsB          TTVKNLPVKAKRRRYADGLVKFLAGKSSMFSYLYSIMIVRSGDLGSLMIRLLIVGMLIVA 292
M5_EcsB          TTVKNLPVKAKRRRYADGLVKFLAGKSSMFSYLYSIMIVRSGDLGSLMIRLLIVGMLIVA 292
TM22_EcsB        TTVKNLPVKAKRRRYADGLVKFLAGKSSMFSYLYSIMIVRSGDLGSLMIRLLIVGMLIVA 292
M4_EcsB          TTVKNLPVKAKRRRYADGLVKFLAGKSSMFSYLYSIMIVRSGDLGSLMIRLLIVGMLIVA 292
M16_EcsB         TTVKNLPVKAKRRRYADGLVKFLAGKSSMFSYLYSIMIVRSGDLGSLMIRLLIVGMLIVA 292
M8_EcsB          YHREKFAR----------------------------------------------------- 289
M12_EcsB         ------------------------------------------------------------
M19_EcsB         ------------------------------------------------------------
M11_EcsB         ------------------------------------------------------------
M7_EcsB          ------------------------------------------------------------

EcsB_2002wt      YSDQEWLKLMVGALTTYLIVIQLVDVYRPVQANSMVQVYPVDAEMAQKGLLSVLKRII 350
M1_EcsB          YSDQEWLKLMVGALTTYLIVIQLVDVYRPVQANSMVQVYPVDAEMAQKGLLSVLKRII 350
M5_EcsB          YSDQEWLKLMVGALTTYLIVIQLVDVYRPVQANSMVQVYPVDAEMAQKGLLSVLKRII 350
TM22_EcsB        YSDQEWLKLMVGALTTYLIVIQLVDVYRPVQANSMVQVYPVDAEMAQKGLLSVLKRII 350
M4_EcsB          YSDQEWLKLMVGALTTYLIVIQLVDVYRPVQANSMVQVYPVDAEMAQKGLLSVLKRII 350
M16_EcsB         YSDQEWLKLMVGALTTYLIVIQLVDVYRPVQANSMVQVYPVDAEMAQKGLLSVLKRII 350
M8_EcsB          ------------------------------------------------------------
```

```
M12_EcsB        ------------------------------------------------------------
M19_EcsB        ------------------------------------------------------------
M11_EcsB        ------------------------------------------------------------
M7_EcsB         ------------------------------------------------------------

EcsB_2002wt     LLNAILLTLAGCVTNWSLSFLGWAIVVQAITAWLLTQWLVPCYIKKINRX 400
M1_EcsB         LLNAILLTLAGCVTNWSLSFLGWAIVVQAITAWLLTQWLVPCYIKKIKRK 400
M5_EcsB         LLNAILLTLAGCVTNWSLSFLGWAIVVQAITAWLLTQWLVPCYIKKINRX 400
TM22_EcsB       LLNAILLTLAGCVTNWSLSFLGWAIVVQAITAWLLTQWLVPCYIKKINRX 400
M4_EcsB         LYKRDFINAGRLRD---------------------------------- 364
M16_EcsB        LYKRDFINAGRLRD---------------------------------- 364
M8_EcsB         ------------------------------------------------
M12_EcsB        ------------------------------------------------
M19_EcsB        ------------------------------------------------
M11_EcsB        ------------------------------------------------
M7_EcsB         ------------------------------------------------
```

A19: MSA of EcsB 3358

```
3358wt_EcsB     MGEIFSQRLSRHFKKMSKYLRYILNDHFVLVCMFLLGGLGLYYSQLLKELPRDFVWGRPL 60
E24_EcsB        MGEIFSQRLSRHFKKMSKYLRYILNDHFVLVCMFLLGGLGLYYSQLLKELPRDFVWGRPL 60
                ************************************************************

3358wt_EcsB     ILLGWLLLIQVGKLATLTEEPDKVFLLPKEKQFAAYLKRALRYSLLLPIVVSFLGSGLLM 120
E24_EcsB        ILLGWLLLIQVGKLATLTEEPDKVFLLPKEKQFAAYLKRALRYSLLLPIVVSFLGSGLLM 120
                ************************************************************

3358wt_EcsB     PLIVVTTGWSFQTFFLFLVMLVCMIYTHLSLQSYGLYHLSSTTYRSWWFVWLISSLLIMT 180
E24_EcsB        PLIVVTTGWSFQTFFLFLVMLVCMIYTHLSLQSYGLYHLSSTTYRSWWFVWLISSLLIMT 180
                ************************************************************

3358wt_EcsB     GAIYWTPWVGVIGGIILAVCLSSIWQNKMKKSFLDWEKMIQKEQNRMHRIYKFIQLFTDI 240
E24_EcsB        GAIYWTPWVGVIGGIILAVCLSSIWQNKMKKSFLDWEKMIQKEQNRMHRIYKFIQLFTDI 240
                ************************************************************

3358wt_EcsB     PEVSSTVKRRKYLDPLLGVVKKTSENTYAYLFIRSFLRGSEYSGLLFRLILVGGVLLFFL 300
E24_EcsB        PEVSSTVKRRKYLDPLLGVVKKNI---------------------------------- 264
                **********************.

3358wt_EcsB     QEFWIALVVALLFVYLIGFQLIPMYTQFDYMVMTQLYPISIEKKQAAIRRLISGALSVAA 360
E24_EcsB        ------------------------------------------------------------

3358wt_EcsB     IIFGGIVCLRLGSLQNSLIILGALIVEVVIFTKMYVPMRLNKL 403
E24_EcsB        -------------------------------------------
```

A20: MSA of EcsB 1546

```
1546 EcsB       MLMNSLFEERRKVYYRQNLKYLRYVFNDHFVLFLMILLGALAVQYAQFLQAHSLNTAGKV 60
G18_EcsB        MLMNSLFEERRKVYYRQNLKYLRYVFNDHFVLFLMILLGALAVQYAQFLQAHSLNTAGKV 60
G13_EcsB        MLMNSLFEERRKVYYRQNLKYLRYVFNDHFVLFLMILLGALAVQYA-------------- 46
                ********************************************

1546 EcsB       GLVLLITILSQIFGRLASFVEPADKVFLLPQEKAVRKHLLLSCLRSLLFPALISLVLVGI 120
G18_EcsB        GLVLLITILSQIFGRLASFVEPADKVFLLPQEKAVRKHLLLSCLRSLLFPALISLVLVGI 120
G13 EcsB        ------------------------------------------------------------

1546_EcsB       VAPLLKWSFFYLLLWFILLVLLKAAGLGLRLRQLMPNGIISWERLIAYEENRKTNTLRFF 180
G18_EcsB        VAPLLK---------------------------------------------------- 126
G13 EcsB        ------------------------------------------------------------

1546_EcsB       ALFTNVKGLKTQSRRRKYLDFLLPKTLRTYEYLFSRAFLRSGDYLSLTLRLIALSVLSLI 240
G18 EcsB        ------------------------------------------------------------
G13_EcsB        ------------------------------------------------------------

1546_EcsB       FIGNPILAVILASVFNYLLLFQLFALQDSFDYQVLTRIYPLRQSSKFAAIKNVLSRIMLF 300
G18 EcsB        ------------------------------------------------------------
G13 EcsB        ------------------------------------------------------------

1546_EcsB       ATVVEVIFGLVFLQPRLYLIVLLLVNFLLVKFYIKMRLKGK 341
G18_EcsB        -----------------------------------------
G13_EcsB        -----------------------------------------
```

A21: The complete *E. faecalis 3358* CLC workbench probabilistic variant detection results.

solate	Mapping	Reference Position	Consensus Position	Variant type	Reference	Allele variants	Frequencies	Coverage	Overlapping annotations
E2	Contig_16 mapping	56992	56992	InDel	A	-/A	63.9/36.1	36	
	Contig_16 mapping	57009	57008	SNV	T	T/C	71.1/28.9	38	
	Contig_5 mapping	141178	141141	SNV	C	A	100	33	CDS: ecsA_2
E3	Contig_13 mapping	28928	28923	SNV	A	C	97.4	39	CDS: PROKKA_00689
	Contig_16 mapping	56992	56992	InDel	A	-/A	67.7/32.3	62	
	Contig_2 mapping	5691	5691	InDel	CAAAAAAT	-	93.2	74	CDS: rasP
	Contig_21 mapping	58	58	InDel	A	-	100	8	
	Contig_3 mapping	233483	233483	SNV	T	T/C	77.0/23.0	456	CDS: PROKKA_01775
	Contig_9 mapping	59315	59235	SNV	G	G/A	74.7/25.3	79	CDS: PROKKA_02751
	Contig_9 mapping	59320	59240	SNV	A	A/G	75.3/23.5	81	CDS: PROKKA_02751
E4	Contig_16 mapping	56992	56992	InDel	A	-/A	68.6/31.4	35	
	Contig_21 mapping	58	58	InDel	A	-	100	6	
	Contig_4 mapping	144738	144738	InDel	T	T/-	77.8/22.2	9	CDS: yheI_2
E8	Contig_16 mapping	56992	56992	InDel	A	-/A	60.5/39.5	81	
	Contig_5 mapping	141090	141053	InDel	A	-/A	55.6/44.4	18	CDS: ecsA_2
	Contig_9 mapping	59320	59240	SNV	A	A/G	69.4/30.6	62	CDS: PROKKA_02751
E11	Contig_16 mapping	56992	56974	InDel	A	-/A	53.3/46.7	30	
	Contig_3 mapping	233483	233483	SNV	T	T/C	71.1/28.3	173	CDS: PROKKA_01775
	Contig_5 mapping	141178	141141	SNV	C	A	100	23	CDS: ecsA_2
E12	Contig_15 mapping	828	828	InDel	-	-/T	69.6/30.4	23	
	Contig_16 mapping	56992	56992	InDel	A	-/A	52.9/47.1	34	
	Contig_3 mapping	233483	233483	SNV	T	T/C	76.2/23.8	168	CDS: PROKKA_01775
	Contig_5 mapping	141178	141141	SNV	C	A	100	17	CDS: ecsA_2
E14	Contig_15 mapping	828	828	InDel	-	-/T	74.6/25.4	67	
	Contig_15 mapping	28172	28172	SNV	C	A/C	66.7/33.3	60	CDS: yjlD
	Contig_16 mapping	56992	56992	InDel	A	-/A	67.6/32.4	74	
	Contig_21 mapping	58	58	InDel	A	-	100	10	
	Contig_5 mapping	141178	141141	SNV	C	A	100	73	CDS: ecsA_2
	Contig_9 mapping	59320	59240	SNV	A	A/G	72.5/27.5	120	CDS: PROKKA_02751
E24	Contig_1 mapping	168503	168503	SNV	C	A	98.4	62	CDS: bglH_1
	Contig_10 mapping	26975	26975	SNV	A	G/A	57.1/41.4	70	CDS: PROKKA_00379
	Contig_10 mapping	26978	26978	SNV	A	G/A	55.4/44.6	65	CDS: PROKKA_00379
	Contig_10 mapping	26981	26981	SNV	A	T/A	53.2/46.8	62	CDS: PROKKA_00379
	Contig_10 mapping	26985	26985	SNV	C	T/C	50.8/49.2	63	CDS: PROKKA_00379
	Contig_10 mapping	26987	26987	SNV	A	G/A	50.8/49.2	63	CDS: PROKKA_00379
	Contig_10 mapping	27003	27003	SNV	A	A/G	60.0/40.0	55	CDS: PROKKA_00379
	Contig_10 mapping	27005	27005	SNV	A	A/G	58.2/41.8	55	CDS: PROKKA_00379
	Contig_10 mapping	27011	27011	SNV	G	G/A	63.0/37.0	54	CDS: PROKKA_00379
	Contig_10 mapping	27015	27015	SNV	G	G/C	70.0/30.0	50	CDS: PROKKA_00379
	Contig_10 mapping	27017	27017	SNV	A	A/G	70.0/30.0	50	CDS: PROKKA_00379
	Contig_10 mapping	27023	27023	SNV	C	C/T	70.0/30.0	50	CDS: PROKKA_00379
	Contig_10 mapping	27028	27028	MNV	GA	GA/AG	75.9/24.1	87	CDS: PROKKA_00379
	Contig_10 mapping	29277	29277	MNV	CT	CT/TC	75.0/25.0	48	CDS: PROKKA_00381
	Contig_10 mapping	29283	29283	SNV	A	A/G	72.4/27.6	29	CDS: PROKKA_00381
	Contig_10 mapping	29289	29289	SNV	C	C/T	72.4/27.6	29	CDS: PROKKA_00381
	Contig_10 mapping	29291	29291	SNV	G	G/C	73.3/26.7	30	CDS: PROKKA_00381
	Contig_10 mapping	29295	29295	SNV	T	T/C	71.0/29.0	31	CDS: PROKKA_00381
	Contig_10 mapping	29301	29301	SNV	C	C/T	64.3/35.7	28	CDS: PROKKA_00381
	Contig_10 mapping	29303	29303	SNV	C	C/T	64.3/35.7	28	CDS: PROKKA_00381
	Contig_10 mapping	29319	29319	SNV	C	C/T	56.7/43.3	30	CDS: PROKKA_00381
	Contig_10 mapping	29321	29321	SNV	A	A/G	56.7/43.3	30	CDS: PROKKA_00381
	Contig_10 mapping	29325	29325	SNV	A	A/T	53.6/46.4	28	CDS: PROKKA_00381
	Contig_10 mapping	29328	29328	SNV	C	T/C	51.6/48.4	31	CDS: PROKKA_00381
	Contig_10 mapping	29331	29331	SNV	C	T/C	53.1/46.9	32	CDS: PROKKA_00381
	Contig_13 mapping	67706	67701	SNV	A	G	100	57	CDS: luxO_1
	Contig_13 mapping	67973	67968	SNV	G	T	100	65	CDS: luxO_1
	Contig_16 mapping	56992	56992	InDel	A	-/A	64.9/35.1	77	
	Contig_2 mapping	100859	100859	SNV	A	C	100	43	CDS: PROKKA_01228

	Contig_2 mapping	108105	108105	SNV	G	A	100	69	CDS: mepM_1
	Contig_2 mapping	271471	271469	SNV	T	C	100	31	CDS: hisS
	Contig_21 mapping	58	58	InDel	A	-	100	11	
	Contig_22 mapping	9074	9074	SNV	A	C	98.5	67	CDS: PROKKA_01482
	Contig_3 mapping	209677	209677	SNV	G	T	100	43	CDS: PROKKA_01754
	Contig_3 mapping	221921	221921	SNV	C	T	100	68	CDS: PROKKA_01764
	Contig_3 mapping	233483	233483	SNV	T	T/C	74.3/25.2	444	CDS: PROKKA_01775
	Contig_4 mapping	170830	170830	SNV	T	C	100	19	
	Contig_5 mapping	140041	140004	InDel	-	T	98.6	70	CDS:ecsB
	Contig_9 mapping	59315	59287	SNV	G	G/A	71.4/27.6	98	CDS: PROKKA_02751
	Contig_9 mapping	59320	59292	SNV	A	A/G	67.6/31.4	105	CDS: PROKKA_02751
E25	Contig_1 mapping	168503	168503	SNV	C	A	98.7	77	CDS: bglH_1
	Contig_13 mapping	67706	67701	SNV	A	G	100	56	CDS: luxO_1
	Contig_13 mapping	67973	67968	SNV	G	T	100	82	CDS: luxO_1
	Contig_16 mapping	8954	8954	InDel	C	-	100	8	
	Contig_16 mapping	8955	8954	InDel	G	-	100	8	
	Contig_16 mapping	8956	8954	InDel	T	-	100	8	
	Contig_16 mapping	8957	8954	InDel	T	-	100	8	
	Contig_16 mapping	8958	8954	InDel	T	-	100	8	
	Contig_16 mapping	8959	8954	InDel	T	-	100	8	
	Contig_16 mapping	8960	8954	InDel	A	-	100	8	
	Contig_16 mapping	8961	8954	InDel	C	-	100	8	
	Contig_16 mapping	8962	8954	InDel	T	-	100	8	
	Contig_16 mapping	8963	8954	InDel	A	-	100	8	
	Contig_16 mapping	8964	8954	InDel	A	-	100	8	
	Contig_16 mapping	8965	8954	InDel	A	-	100	8	
	Contig_16 mapping	8966	8954	InDel	A	-	100	8	
	Contig_16 mapping	8967	8954	InDel	C	-	100	8	
	Contig_16 mapping	8968	8954	InDel	G	-	100	8	
	Contig_16 mapping	8969	8954	InDel	T	-	100	8	
	Contig_16 mapping	8970	8954	InDel	T	-	100	8	
	Contig_16 mapping	8971	8954	InDel	T	-	100	8	
	Contig_16 mapping	8972	8954	InDel	T	-	100	8	
	Contig_16 mapping	8973	8954	InDel	A	-	100	8	
	Contig_16 mapping	8974	8954	InDel	A	-	100	8	
	Contig_16 mapping	8975	8954	InDel	T	-	100	8	
	Contig_16 mapping	8976	8954	InDel	T	-	100	8	
	Contig_16 mapping	8977	8954	InDel	A	-	100	8	
	Contig_16 mapping	8978	8954	InDel	G	G/-	70.0/30.0	20	
	Contig_16 mapping	8979	8955	InDel	A	A/-	70.0/30.0	20	
	Contig_16 mapping	8980	8956	InDel	A	A/-	70.0/30.0	20	
	Contig_16 mapping	8981	8957	InDel	T	T/-	70.0/30.0	20	
	Contig_16 mapping	8982	8958	MNV	G	T/-	70.0/30.0	20	
	Contig_16 mapping	8983	8959	InDel	A	A/-	70.0/30.0	20	
	Contig_16 mapping	8984	8960	InDel	A	A/-	70.0/30.0	20	
	Contig_16 mapping	8985	8961	InDel	A	A/-	70.0/30.0	20	
	Contig_16 mapping	8986	8962	InDel	A	A/-	70.0/30.0	20	
	Contig_16 mapping	56992	56968	InDel	A	-/A	65.5/34.5	116	
	Contig_16 mapping	57009	56984	SNV	T	T/C	77.9/22.1	113	
	Contig_2 mapping	5699	5699	InDel	-	CAAAAAAT	83.8	80	CDS: rasP
	Contig_2 mapping	100859	100867	SNV	A	C	98.5	67	CDS: PROKKA_01228
	Contig_2 mapping	108105	108113	SNV	G	A	100	86	CDS: mepM_1
	Contig_2 mapping	271471	271477	SNV	T	C	100	50	CDS: hisS
	Contig_21 mapping	58	58	InDel	A	-	100	12	
	Contig_22 mapping	9074	9074	SNV	A	C	98.9	90	CDS: PROKKA_01482
	Contig_3 mapping	209677	209677	SNV	G	T	98.3	60	CDS: PROKKA_01754
	Contig_3 mapping	221921	221921	SNV	C	T	100	70	CDS: PROKKA_01764
	Contig_3 mapping	233483	233483	SNV	T	T/C	73.0/26.8	596	CDS: PROKKA_01775
	Contig_4 mapping	170830	170830	SNV	T	C	100	27	
	Contig_9 mapping	59320	59240	SNV	A	A/G	72.9/27.1	140	CDS: PROKKA_02751
E26	Contig_10 mapping	26975	26975	SNV	A	A/G	56.3/43.8	64	CDS: PROKKA_00379
	Contig_10 mapping	26978	26978	SNV	A	A/G	59.0/41.0	61	CDS: PROKKA_00379
	Contig_10 mapping	26981	26981	SNV	A	A/T	60.0/40.0	60	CDS: PROKKA_00379

	Contig_10 mapping	26985	26985	SNV	C	C/T	64.1/35.9	64	CDS: PROKKA_00379
	Contig_10 mapping	26987	26987	SNV	A	A/G	64.1/35.9	64	CDS: PROKKA_00379
	Contig_10 mapping	27003	27003	SNV	A	A/G	68.3/31.7	60	CDS: PROKKA_00379
	Contig_10 mapping	27005	27005	SNV	A	A/G	68.9/31.1	61	CDS: PROKKA_00379
	Contig_10 mapping	29319	29319	SNV	C	C/T	62.5/37.5	48	CDS: PROKKA_00381
	Contig_10 mapping	29321	29321	SNV	A	A/G	62.5/37.5	48	CDS: PROKKA_00381
	Contig_10 mapping	29325	29325	SNV	A	A/T	58.7/41.3	46	CDS: PROKKA_00381
	Contig_10 mapping	29328	29328	SNV	C	C/T	56.3/43.8	48	CDS: PROKKA_00381
	Contig_10 mapping	29331	29331	SNV	C	C/T	54.0/46.0	50	CDS: PROKKA_00381
	Contig_16 mapping	56992	56992	InDel	A	-/A	72.2/27.8	97	
	Contig_21 mapping	58	58	InDel	A	-	100	10	
	Contig_5 mapping	141178	141141	SNV	C	A	100	46	CDS: ecsA_2
	Contig_9 mapping	59320	59240	SNV	A	A/G	69.7/30.3	99	CDS: PROKKA_02751

A22: The complete *L. garvieae* 1546 CLC workbench probabilistic variant detection results.

Isolate	Mapping	Reference Position	Consensus Position	Variant type	Reference	Allele variants	Frequencies	Coverage	Overlapping annotations
G13	Contig_19 mapping	7663	7663	SNV	C	C/A	73.6/26.4	201	CDS: PROKKA_01037
	Contig_19 mapping	12244	12244	SNV	G	A	92.3	13	
	Contig_2 mapping	105252	105252	SNV	C	T	100	31	CDS: ecsB
	Contig_20 mapping	11725	11653	InDel	-	-/T	75.0/25.0	8	
	Contig_21 mapping	523	523	SNV	A	A/C	73.2/26.1	153	
	Contig_8 mapping	91239	91239	InDel	G	-	79.5	39	CDS: PROKKA_02168
G11	Contig_19 mapping	7663	7663	SNV	C	C/A	73.6/26.4	246	CDS: PROKKA_01037
	Contig_19 mapping	12244	12244	SNV	G	A	100	18	
	Contig_2 mapping	104573	104573	SNV	G	T	100	46	CDS: ecsA_1
	Contig_21 mapping	523	523	SNV	A	A/C	78.1/21.9	269	
	Contig_8 mapping	91239	91239	InDel	G	-	94.8	77	CDS: PROKKA_02168
	Contig_9 mapping	41078	41078	InDel	A	-/A	83.3/16.7	12	
G12	Contig_17 mapping	9109	9109	SNV	C	T	100	51	CDS: rasP
	Contig_19 mapping	7663	7663	SNV	C	C/A	72.5/27.5	255	CDS: PROKKA_01037
	Contig_19 mapping	12244	12244	SNV	G	A	95.2	21	
	Contig_21 mapping	523	523	SNV	A	A/C	77.0/22.8	391	
	Contig_8 mapping	91239	91239	InDel	G	-/G	75.0/25.0	52	CDS: PROKKA_02168
	Contig_9 mapping	41078	41078	InDel	A	-/A	70.0/25.0	20	
G20	Contig_17 mapping	8780	8780	SNV	C	A	100	57	CDS: rasP
	Contig_19 mapping	7663	7663	SNV	C	C/A	76.4/23.6	284	CDS: PROKKA_01037
	Contig_19 mapping	12244	12244	SNV	G	A	100	17	
	Contig_21 mapping	523	523	SNV	A	A/C	75.5/24.5	339	
	Contig_8 mapping	91239	91239	InDel	G	-/G	72.1/26.5	68	CDS: PROKKA_02168
	Contig_9 mapping	41078	41078	InDel	A	-	100	10	
G19	Contig_12 mapping	53136	53136	SNV	G	C	99.2	126	CDS: PROKKA_00822
	Contig_19 mapping	7663	7663	SNV	C	C/A	67.0/33.0	330	CDS: PROKKA_01037
	Contig_19 mapping	12244	12244	SNV	G	A	100	23	
	Contig_2 mapping	13642	13642	InDel	T	T/-	77.8/22.2	18	
	Contig_2 mapping	104874	104874	SNV	G	T	100	95	CDS: ecsA_1
	Contig_20 mapping	78	15	InDel	-	A	100	10	CDS: PROKKA_01273
	Contig_21 mapping	523	523	SNV	A	A/C	75.7/23.6	428	
	Contig_3 mapping	124816	124816	InDel	A	A/-	72.7/27.3	11	
	Contig_8 mapping	91239	91239	InDel	G	-	83.6	61	CDS: PROKKA_02168
G18	Contig_19 mapping	7663	7663	SNV	C	C/A	73.5/26.5	347	CDS: PROKKA_01037
	Contig_19 mapping	12244	12244	SNV	G	A	100	25	
	Contig_19 mapping	12276	12276	SNV	C	A	100	10	
	Contig_2 mapping	105493	105493	SNV	G	A	100	55	CDS: ecsB
	Contig_20 mapping	257	185	SNV	A	A/C	64.8/35.2	71	CDS: PROKKA_01273
	Contig_20 mapping	11725	11653	InDel	-	-/T	75.0/25.0	8	
	Contig_21 mapping	523	523	SNV	A	A/C	72.9/27.1	505	
	Contig_8 mapping	91239	91239	InDel	G	-	86.9	84	CDS: PROKKA_02168
	Contig_9 mapping	41078	41078	InDel	A	-/A	50.0/50.0	16	
G14	Contig_1 mapping	361233	361233	SNV	G	T	100	84	CDS: PROKKA_00385

	Contig_19 mapping	7663	7663	SNV	C	C/A	68.6/31.4	236	CDS: PROKKA_01037
	Contig_19 mapping	12244	12244	SNV	G	A	100	21	
	Contig_2 mapping	104851	104851	SNV	C	A	100	55	CDS: ecsA_1
	Contig_2 mapping	225104	225104	SNV	A	G	18.2	11	
	Contig_21 mapping	523	523	SNV	A	A/C	77.9/21.5	298	
	Contig_3 mapping	124816	124816	InDel	A	A/-	75.0/25.0	16	
	Contig_8 mapping	91239	91239	InDel	G	-/G	74.4/25.6	39	CDS: PROKKA_02168
	Contig_9 mapping	41077	41077	MNV	AA	A/-	52.5/32.5	40	
G15	Contig_19 mapping	7663	7663	SNV	C	C/A	77.0/23.0	461	CDS: PROKKA_01037
	Contig_19 mapping	12244	12244	SNV	G	A	100	18	
	Contig_2 mapping	104558	104558	SNV	C	T	100	60	CDS: ecsA_1
	Contig_21 mapping	523	523	SNV	A	A/C	76.8/23.2	315	
	Contig_8 mapping	91239	91239	InDel	G	-	82.3	62	CDS: PROKKA_02168
	Contig_9 mapping	41078	41078	InDel	A	-	78.6	14	
G17	Contig_17 mapping	8524	8524	SNV	C	T	100	112	CDS: rasP
	Contig_19 mapping	7663	7663	SNV	C	C/A	74.9/24.9	366	CDS: PROKKA_01037
	Contig_19 mapping	12244	12244	SNV	G	A	100	24	
	Contig_19 mapping	12276	12276	SNV	C	A	100	11	
	Contig_21 mapping	523	523	SNV	A	A/C	71.7/28.3	459	
	Contig_3 mapping	222261	222261	SNV	C	T	98.2	55	CDS: PROKKA_01566
	Contig_8 mapping	91239	91239	InDel	G	-	81.3	75	CDS: PROKKA_02168
	Contig_9 mapping	41078	41078	InDel	A	-	93.8	16	
G16	Contig_19 mapping	7663	7663	SNV	C	C/A	77.4/22.6	438	CDS: PROKKA_01037
	Contig_19 mapping	12244	12244	SNV	G	A	96.2	26	
	Contig_2 mapping	104749	104749	SNV	C	A	100	93	CDS: ecsA_1
	Contig_21 mapping	523	523	SNV	A	A/C	74.9/24.9	358	
	Contig_3 mapping	124817	124817	InDel	-	-/A	73.3/26.7	15	
	Contig_8 mapping	91239	91239	InDel	G	-	82.6	86	CDS: PROKKA_02168
	Contig_9 mapping	41077	41076	MNV	AA	A/AA/-	45.5/27.3/27.3	22	

A23: The complete *P. acidilactici* 2002 CLC workbench probabilistic variant detection results.

Isolate	Mapping	Reference Position	Consensus Position	Variant type	Reference	Allele variants	Frequencies	Coverage	Overlapping annotations
M5	contig_5	1110	1110	SNV	G	G/A	79.3/20.7	140	CDS: Mobile element protein
	contig_5	1231	1231	SNV	A	A/G	57.1/42.9	196	CDS: Mobile element protein
	contig_5	1314	1314	SNV	G	G/A	73.2/26.4	235	
	contig_5	1340	1340	SNV	G	G/T	51.0/49.0	243	
	contig_5	1363	1363	SNV	A	A/G	51.1/48.9	237	CDS: Mobile element protein
	contig_5	1387	1387	SNV	T	T/C	52.4/47.2	231	CDS: Mobile element protein
	contig_5	1389	1389	SNV	T	T/C	77.1/22.9	236	CDS: Mobile element protein
	contig_5	1434	1434	SNV	T	T/C	51.6/47.9	215	CDS: Mobile element protein
	contig_5	1442	1442	SNV	A	A/C	51.1/48.9	223	CDS: Mobile element protein
	contig_5	1486	1486	SNV	C	C/T	76.8/22.7	220	CDS: Mobile element protein
	contig_5	1534	1534	SNV	G	G/A	54.6/45.4	227	CDS: Mobile element protein
	contig_5	1537	1537	SNV	T	T/C	53.3/45.4	229	CDS: Mobile element protein
	contig_5	1716	1716	SNV	T	T/C	56.9/42.7	248	CDS: Mobile element protein
	contig_5	1768	1768	SNV	G	G/C	61.2/38.8	219	CDS: Mobile element protein
	contig_5	1777	1777	SNV	A	A/C	62.1/37.9	206	CDS: Mobile element protein
	contig_5	32344	32344	InDel	-	CGAACGGTTTTTGTTCGTT	92.6	54	
	contig_5	47690	47709	SNV	G	A	100	68	CDS: Mobile element protein
	contig_6	170837	170837	InDel	-	GCCAAACTTAGGCAT	95.7	46	
	contig_6	343959	343974	SNV	G	A	100	92	CDS: Metalloprotease RseP
	contig_1	182726	182726	InDel	-	ATT	100	105	
	contig_8	72413	72413	SNV	A	T	100	71	
	contig_11	21	21	SNV	A	A/G	50.0/50.0	24	
	contig_12	7400	7400	SNV	C	T	100	65	
	contig_13	7487	7487	SNV	C	T	100	67	
	contig_14	1100	1100	SNV	A	A/G	56.5/43.5	161	
	contig_14	1292	1292	SNV	C	T/C	58.8/41.2	119	

M1	contig_5	1110	1110	SNV	G	G/A	76.6/23.4	145	CDS: Mobile element protein
	contig_5	1231	1231	SNV	A	G/A	54.6/44.4	196	CDS: Mobile element protein
	contig_5	1314	1314	SNV	G	G/A	68.5/31.1	235	
	contig_5	1340	1340	SNV	G	T/G	61.8/38.2	238	
	contig_5	1363	1363	SNV	A	G/A	61.3/38.7	238	CDS: Mobile element protein
	contig_5	1387	1387	SNV	T	C/T	60.2/39.3	244	CDS: Mobile element protein
	contig_5	1389	1389	SNV	T	T/C	71.5/28.0	246	CDS: Mobile element protein
	contig_5	1434	1434	SNV	T	C/T	59.9/40.1	227	CDS: Mobile element protein
	contig_5	1442	1442	SNV	A	C/A	58.7/41.3	235	CDS: Mobile element protein
	contig_5	1486	1486	SNV	C	C/T	69.8/30.2	235	CDS: Mobile element protein
	contig_5	1534	1534	SNV	G	A/G	59.3/40.7	246	CDS: Mobile element protein
	contig_5	1537	1537	SNV	T	C/T	59.5/40.5	242	CDS: Mobile element protein
	contig_5	1716	1716	SNV	T	C/T	53.8/46.2	251	CDS: Mobile element protein
	contig_5	1768	1768	SNV	G	C/G	50.2/49.8	219	CDS: Mobile element protein
	contig_5	1777	1777	SNV	A	A/C	51.7/48.3	207	CDS: Mobile element protein
	contig_5	1796	1796	SNV	T	T/G	77.2/22.8	197	CDS: Mobile element protein
	contig_5	1798	1798	MNV	AC	AC/CA	76.2/23.8	332	CDS: Mobile element protein
	contig_5	32344	32344	InDel	-	CGAACGGTTTTTGTTCGTT	95.7	46	
	contig_5	47690	47709	SNV	G	A	100	63	CDS: Mobile element protein
	contig_2	67916	67916	SNV	G	A	98.4	64	
	contig_6	170837	170837	InDel	-	GCCAAACTTAGGCAT	92.1	63	
	contig_6	344153	344168	InDel	T	-	100	81	CDS: Metalloprotease RseP
	contig_1	1304	1304	SNV	T	T/A	70.3/27.0	37	CDS: hypothetical protein
	contig_1	182726	182726	InDel	-	ATT	99.2	118	
	contig_8	72413	72413	SNV	A	T	100	59	
	contig_11	21	21	SNV	A	A/G	60.9/39.1	23	
	contig_12	19	19	SNV	T	T/G	78.9/21.1	19	
	contig_12	7400	7400	SNV	C	T	100	61	
	contig_12	7487	7487	SNV	C	T	100	61	
	contig_14	1100	1100	SNV	A	A/G	57.1/42.9	147	
	contig_14	1292	1292	SNV	C	T/C	50.0/47.1	104	
M16	contig_5	1110	1110	SNV	G	G/A	72.6/27.4	135	CDS: Mobile element protein
	contig_5	1231	1231	SNV	A	G/A	59.8/40.2	224	CDS: Mobile element protein
	contig_5	1314	1314	SNV	G	G/A	68.8/31.2	237	
	contig_5	1340	1340	SNV	G	T/G	66.8/33.2	247	
	contig_5	1363	1363	SNV	A	G/A	66.5/33.5	233	CDS: Mobile element protein
	contig_5	1387	1387	SNV	T	C/T	64.8/35.2	236	CDS: Mobile element protein
	contig_5	1389	1389	SNV	T	T/C	67.1/32.9	237	CDS: Mobile element protein
	contig_5	1434	1434	SNV	T	C/T	66.8/32.3	226	CDS: Mobile element protein
	contig_5	1442	1442	SNV	A	C/A	67.1/32.9	237	CDS: Mobile element protein
	contig_5	1486	1486	SNV	C	C/T	68.7/30.9	249	CDS: Mobile element protein
	contig_5	1534	1534	SNV	G	A/G	64.5/35.5	234	CDS: Mobile element protein
	contig_5	1537	1537	SNV	T	C/T	63.8/36.2	232	CDS: Mobile element protein
	contig_5	1716	1716	SNV	T	C/T	60.2/39.8	259	CDS: Mobile element protein
	contig_5	1768	1768	SNV	G	C/G	56.0/44.0	232	CDS: Mobile element protein
	contig_5	1777	1777	SNV	A	C/A	53.2/46.8	218	CDS: Mobile element protein
	contig_5	1796	1796	SNV	T	T/G	75.0/25.0	212	CDS: Mobile element protein
	contig_5	1798	1798	MNV	AC	AC/CA	73.7/26.0	354	CDS: Mobile element protein
	contig_5	32344	32344	InDel	-	CGAACGGTTTTTGTTCGTT	97.9	47	
	contig_5	47690	47709	SNV	G	A	100	68	CDS: Mobile element protein
	contig_6	170837	170837	InDel	-	GCCAAACTTAGGCAT	98.3	59	
	contig_6	473899	473914	InDel	-	T	100	54	CDS: ABC transporter EscB
	contig_1	1097	1097	SNV	C	C/T	77.1/20.8	48	CDS: hypothetical protein
	contig_1	1103	1103	SNV	A	A/C	80.0/20.0	45	CDS: hypothetical protein
	contig_1	1259	1259	SNV	G	G/A	75.0/25.0	40	CDS: hypothetical protein
	contig_1	1262	1262	SNV	C	C/G	71.1/26.3	38	CDS: hypothetical protein
	contig_1	1274	1274	SNV	G	G/A	77.1/22.9	35	CDS: hypothetical protein
	contig_1	1280	1280	SNV	A	A/G	81.6/18.4	38	CDS: hypothetical protein
	contig_1	1283	1283	SNV	C	C/A	79.4/20.6	34	CDS: hypothetical protein
	contig_1	1289	1289	SNV	C	C/A	70.6/26.5	34	CDS: hypothetical protein
	contig_1	1292	1292	SNV	G	G/A	76.5/23.5	34	CDS: hypothetical protein

	contig_1	1294	1294	SNV	T	T/C	78.8/21.2	33	CDS: hypothetical protein
	contig_1	1300	1300	MNV	GC	GC/CT	62.2/31.1	45	CDS: hypothetical protein
	contig_1	1304	1304	MNV	TT	TT/A/AT	56.5/26.1/8.7	46	CDS: hypothetical protein
	contig_1	1308	1308	InDel	-	-/A	81.3/18.8	32	CDS: hypothetical protein
	contig_1	1310	1310	SNV	A	A/G	76.5/23.5	34	CDS: hypothetical protein
	contig_1	1346	1346	InDel	-	-/A	75.7/24.3	37	CDS: hypothetical protein
	contig_1	1350	1350	SNV	G	G/A	77.8/22.2	36	CDS: hypothetical protein
	contig_1	1352	1352	InDel	G	G/-	75.0/22.2	36	CDS: hypothetical protein
	contig_1	1354	1354	SNV	T	T/C	77.8/22.2	36	CDS: hypothetical protein
	contig_1	1359	1359	MNV	TCG	TCG/CTA	66.7/29.6	54	CDS: hypothetical protein
	contig_1	1364	1364	SNV	G	G/A	76.5/23.5	34	CDS: hypothetical protein
	contig_1	1367	1367	SNV	A	A/G	76.5/23.5	34	CDS: hypothetical protein
	contig_1	1373	1373	SNV	T	T/C	71.9/25.0	32	CDS: hypothetical protein
	contig_1	1379	1379	SNV	C	C/A	74.2/25.8	31	CDS: hypothetical protein
	contig_1	1385	1385	SNV	T	T/C	69.0/24.1	29	CDS: hypothetical protein
	contig_1	1391	1391	SNV	G	G/C	76.7/23.3	30	CDS: hypothetical protein
	contig_1	1397	1397	SNV	C	C/A	74.2/22.6	31	CDS: hypothetical protein
	contig_1	1412	1412	SNV	C	C/G	79.3/20.7	29	CDS: hypothetical protein
	contig_1	182726	182726	InDel	-	ATT	98.9	94	
	contig_8	72413	72413	SNV	A	T	100	54	
	contig_9	297	297	SNV	T	T/C	76.0/23.6	601	rRNA: rRNA
	contig_11	21	21	SNV	A	G	77.3	22	
	contig_11	77	77	SNV	T	T/G	76.6/23.4	47	
	contig_12	19	19	SNV	T	T/G	56.3/43.8	16	
	contig_12	21	21	MNV	AC	AC/CA	54.8/45.2	31	
	contig_12	166	166	SNV	G	G/A	78.6/21.4	70	
	contig_12	168	168	SNV	A	A/G	77.9/22.1	68	
	contig_12	7400	7400	SNV	C	T	100	62	
	contig_12	7487	7487	SNV	C	T	100	59	
	contig_14	1100	1100	SNV	A	A/G	57.7/41.1	168	
	contig_14	1292	1292	SNV	C	T/C	54.9/45.1	142	
M4	contig_5	1110	1110	SNV	G	G/A	69.0/31.0	113	CDS: Mobile element protein
	contig_5	1231	1231	SNV	A	G/A	66.7/33.3	189	CDS: Mobile element protein
	contig_5	1314	1314	SNV	G	G/A	62.9/37.1	251	
	contig_5	1340	1340	SNV	G	T/G	72.7/27.3	253	
	contig_5	1363	1363	SNV	A	G/A	73.0/27.0	244	CDS: Mobile element protein
	contig_5	1387	1387	SNV	T	C/T	71.9/28.1	235	CDS: Mobile element protein
	contig_5	1389	1389	SNV	T	T/C	67.4/32.6	239	CDS: Mobile element protein
	contig_5	1434	1434	SNV	T	C/T	73.0/27.0	230	CDS: Mobile element protein
	contig_5	1442	1442	SNV	A	C/A	72.8/27.2	239	CDS: Mobile element protein
	contig_5	1486	1486	SNV	C	C/T	67.1/32.9	231	CDS: Mobile element protein
	contig_5	1534	1534	SNV	G	A/G	67.8/32.2	236	CDS: Mobile element protein
	contig_5	1537	1537	SNV	T	C/T	68.3/31.3	230	CDS: Mobile element protein
	contig_5	1716	1716	SNV	T	C/T	64.4/35.6	225	CDS: Mobile element protein
	contig_5	1768	1768	SNV	G	C/G	58.9/41.1	202	CDS: Mobile element protein
	contig_5	1777	1777	SNV	A	C/A	58.3/41.7	199	CDS: Mobile element protein
	contig_5	1796	1796	SNV	T	T/G	74.5/25.5	192	CDS: Mobile element protein
	contig_5	1798	1798	MNV	AC	AC/CA	75.1/24.3	333	CDS: Mobile element protein
	contig_5	32344	32344	InDel	-	CGAACGGTTTTTGTTCGTT	96.4	28	
	contig_5	47690	47709	SNV	G	A	100	63	CDS: Mobile element protein
	contig_5	48208	48227	SNV	T	T/C	71.1/28.9	45	CDS: Mobile element protein
	contig_6	170837	170837	InDel	-	GCCAAACTTAGGCAT	97.8	46	
	contig_6	473899	473914	InDel	-	T	96.9	64	CDS: ABC transporter EscB
	contig_1	1259	1259	SNV	G	G/A	75.0/25.0	48	CDS: hypothetical protein
	contig_1	1262	1262	SNV	C	C/G	73.9/21.7	46	CDS: hypothetical protein
	contig_1	1274	1274	SNV	G	G/A	78.7/21.3	47	CDS: hypothetical protein
	contig_1	1280	1280	SNV	A	A/G	78.0/22.0	50	CDS: hypothetical protein
	contig_1	1283	1283	SNV	C	C/A	78.7/21.3	47	CDS: hypothetical protein
	contig_1	1289	1289	SNV	C	C/A	77.6/22.4	49	CDS: hypothetical protein
	contig_1	1292	1292	SNV	G	G/A	79.2/20.8	48	CDS: hypothetical protein
	contig_1	1294	1294	SNV	T	T/C	78.4/19.6	51	CDS: hypothetical protein

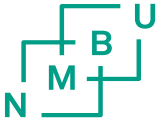
	contig_1	1300	1300	MNV	GC	GC/CT	72.0/26.7	75	CDS: hypothetical protein
	contig_1	1304	1304	MNV	TT	TT/A/AT	62.5/27.8/4.2	72	CDS: hypothetical protein
	contig_1	1308	1308	InDel	-	-/A	79.6/20.4	49	CDS: hypothetical protein
	contig_1	1310	1310	SNV	A	A/G	75.5/22.4	49	CDS: hypothetical protein
	contig_1	1346	1346	InDel	-	-/A	80.4/19.6	56	CDS: hypothetical protein
	contig_1	182726	182726	InDel	-	ATT	98.9	93	
	contig_1	392930	392933	SNV	G	G/A	70.0/30.0	10	
	contig_8	72413	72413	SNV	A	T	100	40	
	contig_9	297	297	SNV	T	T/C	76.6/23.2	698	rRNA: rRNA
	contig_11	21	21	SNV	A	A/G	66.7/33.3	27	
	contig_12	166	166	SNV	G	G/A	75.0/25.0	96	
	contig_12	168	168	SNV	A	A/G	75.5/24.5	94	
	contig_12	7400	7400	SNV	C	T	100	66	
	contig_12	7487	7487	SNV	C	T	100	68	
	contig_14	1100	1100	SNV	A	A/G	51.0/48.3	149	
	contig_14	1292	1292	SNV	C	T/C	56.0/44.0	141	
M12	contig_5	1110	1110	SNV	G	G/A	72.7/27.3	121	CDS: Mobile element protein
	contig_5	1231	1231	SNV	A	G/A	56.0/44.0	166	CDS: Mobile element protein
	contig_5	1314	1314	SNV	G	G/A	71.2/28.8	205	
	contig_5	1340	1340	SNV	G	T/G	63.8/36.2	210	
	contig_5	1363	1363	SNV	A	G/A	61.7/38.3	201	CDS: Mobile element protein
	contig_5	1387	1387	SNV	T	C/T	59.4/40.6	207	CDS: Mobile element protein
	contig_5	1389	1389	SNV	T	T/C	65.1/34.9	209	CDS: Mobile element protein
	contig_5	1434	1434	SNV	T	C/T	61.5/38.5	205	CDS: Mobile element protein
	contig_5	1442	1442	SNV	A	C/A	61.9/38.1	215	CDS: Mobile element protein
	contig_5	1486	1486	SNV	C	C/T	63.0/37.0	208	CDS: Mobile element protein
	contig_5	1534	1534	SNV	G	A/G	60.6/39.4	226	CDS: Mobile element protein
	contig_5	1537	1537	SNV	T	C/T	61.0/38.6	223	CDS: Mobile element protein
	contig_5	1716	1716	SNV	T	C/T	64.9/35.1	205	CDS: Mobile element protein
	contig_5	1768	1768	SNV	G	C/G	53.9/46.1	178	CDS: Mobile element protein
	contig_5	1777	1777	SNV	A	C/A	51.2/48.8	164	CDS: Mobile element protein
	contig_5	32344	32344	InDel	-	CGAACGGTTTTTGTTCGTT	96.4	56	
	contig_5	47690	47709	SNV	G	A	100	68	CDS: Mobile element protein
	contig_6	170837	170837	InDel	-	GCCAAACTTAGGCAT	100	53	
	contig_6	474610	474625	InDel	C	-	100	43	CDS: ABC transporter EscB
	contig_1	1028	1028	SNV	T	T/G	78.4/21.6	37	CDS: hypothetical protein
	contig_1	1304	1304	SNV	T	T/A	78.0/22.0	41	CDS: hypothetical protein
	contig_1	11716	11716	SNV	T	T/G	76.5/23.5	115	CDS: T1SS secreted agglutinin RTX
	contig_1	182726	182726	InDel	-	ATT	98.2	110	
	contig_8	72413	72413	SNV	A	T	100	43	
	contig_11	21	21	SNV	A	A/G	66.7/33.3	15	
	contig_12	19	19	SNV	T	T/G	76.9/23.1	13	
	contig_12	21	21	MNV	AC	AC/CA	72.7/27.3	22	
	contig_12	7400	7400	SNV	C	T	100	57	
	contig_12	7487	7487	SNV	C	T	100	59	
	contig_14	1100	1100	SNV	A	G/A	51.3/48.8	160	
	contig_14	1292	1292	SNV	C	C/T	52.4/47.6	143	
M19	contig_5	1110	1110	SNV	G	G/A	68.2/30.3	66	CDS: Mobile element protein
	contig_5	1231	1231	SNV	A	G/A	70.3/29.7	128	CDS: Mobile element protein
	contig_5	1314	1314	SNV	G	G/A	60.8/39.2	171	
	contig_5	1340	1340	SNV	G	T/G	72.2/27.8	176	
	contig_5	1363	1363	SNV	A	G/A	72.5/27.5	171	CDS: Mobile element protein
	contig_5	1387	1387	SNV	T	C/T	70.0/30.0	170	CDS: Mobile element protein
	contig_5	1389	1389	SNV	T	T/C	68.4/31.6	171	CDS: Mobile element protein
	contig_5	1434	1434	SNV	T	C/T	75.2/24.8	157	CDS: Mobile element protein
	contig_5	1442	1442	SNV	A	C/A	75.8/24.2	161	CDS: Mobile element protein
	contig_5	1486	1486	SNV	C	C/T	57.0/43.0	149	CDS: Mobile element protein
	contig_5	1534	1534	SNV	G	A/G	74.2/25.8	151	CDS: Mobile element protein
	contig_5	1537	1537	SNV	T	C/T	74.5/25.5	149	CDS: Mobile element protein
	contig_5	1716	1716	SNV	T	C/T	71.9/28.1	185	CDS: Mobile element protein
	contig_5	1768	1768	SNV	G	C/G	72.2/27.8	151	CDS: Mobile element protein

	contig_5	1777	1777	SNV	A	C/A	70.4/29.6	135	CDS: Mobile element protein
	contig_5	1796	1796	SNV	T	T/G	63.3/36.7	128	CDS: Mobile element protein
	contig_5	1798	1798	MNV	AC	AC/CA	62.2/37.4	238	CDS: Mobile element protein
	contig_5	32344	32344	InDel	-	CGAACGGTTTTTGTTCGTT	94.1	17	
	contig_5	47690	47709	SNV	G	A	100	40	CDS: Mobile element protein
	contig_5	48208	48227	SNV	T	T/C	60.5/39.5	38	CDS: Mobile element protein
	contig_6	170837	170837	InDel	-	GCCAAACTTAGGCAT	97.1	34	
	contig_6	474847	474862	InDel	C	-	100	22	CDS: ABC transporter EscB
	contig_1	11122	11122	SNV	G	G/A	75.7/23.5	136	
	contig_1	182726	182726	InDel	-	ATT	100	68	
	contig_8	72413	72413	SNV	A	T	100	30	
	contig_9	297	297	SNV	T	T/C	76.4/23.6	589	rRNA: rRNA
	contig_11	21	21	SNV	A	G	81.3	16	
	contig_12	166	166	SNV	G	G/A	77.5/22.5	71	
	contig_12	168	168	SNV	A	A/G	77.1/22.9	70	
	contig_12	7400	7400	SNV	C	T	100	44	
	contig_12	7487	7487	SNV	C	T	100	36	
	contig_14	1100	1100	SNV	A	G/A	52.8/47.2	106	
	contig_14	1292	1292	SNV	C	C/T	51.4/47.6	105	
M11	contig_5	1110	1110	SNV	G	G/A	77.2/22.8	92	CDS: Mobile element protein
	contig_5	1231	1231	SNV	A	G/A	61.1/38.9	149	CDS: Mobile element protein
	contig_5	1314	1314	SNV	G	G/A	69.4/30.6	196	
	contig_5	1340	1340	SNV	G	T/G	63.1/36.9	198	
	contig_5	1363	1363	SNV	A	G/A	62.6/37.4	203	CDS: Mobile element protein
	contig_5	1387	1387	SNV	T	C/T	60.4/39.6	207	CDS: Mobile element protein
	contig_5	1389	1389	SNV	T	T/C	69.6/30.4	207	CDS: Mobile element protein
	contig_5	1434	1434	SNV	T	C/T	62.4/37.6	178	CDS: Mobile element protein
	contig_5	1442	1442	SNV	A	C/A	64.0/36.0	189	CDS: Mobile element protein
	contig_5	1486	1486	SNV	C	C/T	63.7/36.3	193	CDS: Mobile element protein
	contig_5	1534	1534	SNV	G	A/G	67.2/32.8	195	CDS: Mobile element protein
	contig_5	1537	1537	SNV	T	C/T	65.8/34.2	196	CDS: Mobile element protein
	contig_5	1716	1716	SNV	T	C/T	59.5/40.5	195	CDS: Mobile element protein
	contig_5	1768	1768	SNV	G	C/G	54.3/45.7	175	CDS: Mobile element protein
	contig_5	1777	1777	SNV	A	C/A	52.9/46.5	170	CDS: Mobile element protein
	contig_5	1796	1796	SNV	T	T/G	75.3/24.7	166	CDS: Mobile element protein
	contig_5	1798	1798	MNV	AC	AC/CA	75.3/24.7	299	CDS: Mobile element protein
	contig_5	32344	32344	InDel	-	CGAACGGTTTTTGTTCGTT	100	30	
	contig_5	47690	47709	SNV	G	A	98.5	65	CDS: Mobile element protein
	contig_6	170837	170837	InDel	-	GCCAAACTTAGGCAT	100	37	
	contig_6	474847	474862	InDel	C	-	100	26	CDS: ABC transporter EscB
	contig_1	1028	1028	SNV	T	T/G	80.0/20.0	30	CDS: hypothetical protein
	contig_1	1070	1070	InDel	-	-/GGTTG	76.5/23.5	51	CDS: hypothetical protein
	contig_1	1073	1073	SNV	A	A/T	77.8/22.2	27	CDS: hypothetical protein
	contig_1	1097	1097	SNV	C	C/T	76.0/24.0	25	CDS: hypothetical protein
	contig_1	1103	1103	SNV	A	A/C	76.9/23.1	26	CDS: hypothetical protein
	contig_1	1115	1115	SNV	T	T/A	73.9/26.1	23	CDS: hypothetical protein
	contig_1	1125	1125	SNV	C	C/G	73.9/26.1	23	CDS: hypothetical protein
	contig_1	1127	1127	SNV	A	A/C	76.0/24.0	25	CDS: hypothetical protein
	contig_1	1130	1130	SNV	G	G/A	76.0/24.0	25	CDS: hypothetical protein
	contig_1	1132	1132	SNV	T	T/C	76.0/24.0	25	CDS: hypothetical protein
	contig_1	1262	1262	SNV	C	C/G	80.0/20.0	20	CDS: hypothetical protein
	contig_1	1274	1274	SNV	G	G/A	73.9/26.1	23	CDS: hypothetical protein
	contig_1	1280	1280	SNV	A	A/G	75.0/25.0	24	CDS: hypothetical protein
	contig_1	1283	1283	SNV	C	C/A	75.0/25.0	24	CDS: hypothetical protein
	contig_1	1289	1289	SNV	C	C/A	71.4/28.6	21	CDS: hypothetical protein
	contig_1	1292	1292	SNV	G	G/A	70.0/30.0	20	CDS: hypothetical protein
	contig_1	1294	1294	SNV	T	T/C	72.7/27.3	22	CDS: hypothetical protein
	contig_1	1300	1300	MNV	GC	GC/CT	72.1/27.9	43	CDS: hypothetical protein
	contig_1	1304	1304	MNV	TT	TT/A	69.2/30.8	39	CDS: hypothetical protein
	contig_1	1308	1308	InDel	-	-/A	73.9/26.1	23	CDS: hypothetical protein
	contig_1	1310	1310	SNV	A	A/G	70.8/29.2	24	CDS: hypothetical protein

	contig_1	1346	1346	InDel	-	-/A	65.4/34.6	26	CDS: hypothetical protein
	contig_1	1350	1350	SNV	G	G/A	62.5/37.5	24	CDS: hypothetical protein
	contig_1	1352	1352	InDel	G	G/-	59.1/40.9	22	CDS: hypothetical protein
	contig_1	1354	1354	SNV	T	T/C	62.5/37.5	24	CDS: hypothetical protein
	contig_1	1359	1359	MNV	TCG	TCG/CTA	48.7/46.2	39	CDS: hypothetical protein
	contig_1	1364	1364	SNV	G	G/A	60.9/39.1	23	CDS: hypothetical protein
	contig_1	1367	1367	SNV	A	A/G	62.5/37.5	24	CDS: hypothetical protein
	contig_1	1373	1373	SNV	T	T/C	60.9/39.1	23	CDS: hypothetical protein
	contig_1	1379	1379	SNV	C	C/A	61.5/34.6	26	CDS: hypothetical protein
	contig_1	1385	1385	SNV	T	T/C	68.0/32.0	25	CDS: hypothetical protein
	contig_1	1391	1391	SNV	G	G/C	70.4/29.6	27	CDS: hypothetical protein
	contig_1	1397	1397	SNV	C	C/A	73.1/23.1	26	CDS: hypothetical protein
	contig_1	1400	1400	SNV	A	A/G	79.2/20.8	24	CDS: hypothetical protein
	contig_1	1402	1402	SNV	C	C/T	76.2/23.8	21	CDS: hypothetical protein
	contig_1	1409	1409	SNV	G	G/C	78.9/21.1	19	CDS: hypothetical protein
	contig_1	1424	1424	SNV	A	A/G	78.3/21.7	23	CDS: hypothetical protein
	contig_1	1426	1426	MNV	TC	TC/CA	71.4/28.6	42	CDS: hypothetical protein
	contig_1	11716	11716	SNV	T	T/G	72.5/27.5	80	CDS: T1SS secreted agglutinin RTX
	contig_1	11718	11718	SNV	T	T/G	74.7/25.3	83	CDS: T1SS secreted agglutinin RTX
	contig_1	11724	11724	SNV	C	C/A	75.3/24.7	73	CDS: T1SS secreted agglutinin RTX
	contig_1	11728	11728	SNV	A	A/G	73.2/26.8	71	CDS: T1SS secreted agglutinin RTX
	contig_1	11732	11732	SNV	G	G/C	76.0/24.0	75	CDS: T1SS secreted agglutinin RTX
	contig_1	182726	182726	InDel	-	ATT	100	74	
	contig_8	72413	72413	SNV	A	T	100	66	
	contig_11	21	21	SNV	A	A/G	65.2/34.8	23	
	contig_12	7400	7400	SNV	C	T	98	51	
	contig_12	7487	7487	SNV	C	T	98.3	60	
	contig_14	1100	1100	SNV	A	A/G	56.3/43.7	126	
	contig_14	1292	1292	SNV	C	C/T	56.4/43.6	78	
M7	contig_5	1110	1110	SNV	G	G/A	77.8/22.2	117	CDS: Mobile element protein
	contig_5	1231	1231	SNV	A	G/A	53.5/46.5	172	CDS: Mobile element protein
	contig_5	1314	1314	SNV	G	G/A	66.0/34.0	212	
	contig_5	1340	1340	SNV	G	T/G	62.1/37.9	211	
	contig_5	1363	1363	SNV	A	G/A	62.0/38.0	205	CDS: Mobile element protein
	contig_5	1387	1387	SNV	T	C/T	62.8/37.2	199	CDS: Mobile element protein
	contig_5	1389	1389	SNV	T	T/C	71.5/28.5	200	CDS: Mobile element protein
	contig_5	1434	1434	SNV	T	C/T	65.8/34.2	187	CDS: Mobile element protein
	contig_5	1442	1442	SNV	A	C/A	66.7/33.3	195	CDS: Mobile element protein
	contig_5	1486	1486	SNV	C	C/T	71.6/28.4	201	CDS: Mobile element protein
	contig_5	1534	1534	SNV	G	A/G	61.9/38.1	215	CDS: Mobile element protein
	contig_5	1537	1537	SNV	T	C/T	63.0/37.0	211	CDS: Mobile element protein
	contig_5	1716	1716	SNV	T	C/T	57.7/42.3	241	CDS: Mobile element protein
	contig_5	1768	1768	SNV	G	C/G	54.4/45.6	180	CDS: Mobile element protein
	contig_5	1777	1777	SNV	A	C/A	51.2/48.8	172	CDS: Mobile element protein
	contig_5	1796	1796	SNV	T	T/G	75.2/24.2	157	CDS: Mobile element protein
	contig_5	1798	1798	MNV	AC	AC/CA/CC	74.5/24.7/0.4	275	CDS: Mobile element protein
	contig_5	32344	32344	InDel	-	CGAACGGTTTTTGTTCGTT	97.9	47	
	contig_5	47690	47709	SNV	G	A	100	73	CDS: Mobile element protein
	contig_6	170837	170837	InDel	-	GCCAAACTTAGGCAT	96.9	32	
	contig_6	474836	472837	SNV	T	C	100	12	CDS: ABC transporter EscB
	contig_1	182726	182726	InDel	-	ATT	97.7	87	
	contig_8	72413	72413	SNV	A	T	100	56	
	contig_11	21	21	SNV	A	A/G	62.5/37.5	16	
	contig_12	19	19	SNV	T	T/G	72.7/27.3	11	
	contig_12	21	21	MNV	AC	AC/CA	72.7/27.3	22	
	contig_12	7400	7400	SNV	C	T	100	63	
	contig_12	7487	7487	SNV	C	T	100	51	
	contig_14	1100	1100	SNV	A	A/G	55.8/44.2	129	
	contig_14	1292	1292	SNV	C	T/C	55.3/44.7	103	
M22	contig_5	1110	1110	SNV	G	G/A	72.0/28.0	118	CDS: Mobile element protein
	contig_5	1231	1231	SNV	A	G/A	52.5/47.5	202	CDS: Mobile element protein

contig_5	1314	1314	SNV	G	G/A	67.2/32.8	256	
contig_5	1340	1340	SNV	G	T/G	61.4/38.6	249	
contig_5	1363	1363	SNV	A	G/A	61.4/38.6	236	CDS: Mobile element protein
contig_5	1387	1387	SNV	T	C/T	61.3/38.8	240	CDS: Mobile element protein
contig_5	1389	1389	SNV	T	T/C	73.8/26.2	244	CDS: Mobile element protein
contig_5	1434	1434	SNV	T	C/T	62.0/38.0	221	CDS: Mobile element protein
contig_5	1442	1442	SNV	A	C/A	62.3/37.7	228	CDS: Mobile element protein
contig_5	1486	1486	SNV	C	C/T	74.3/25.7	214	CDS: Mobile element protein
contig_5	1534	1534	SNV	G	A/G	61.9/37.7	223	CDS: Mobile element protein
contig_5	1537	1537	SNV	T	C/T	61.4/38.1	223	CDS: Mobile element protein
contig_5	1716	1716	SNV	T	C/T	63.3/36.7	199	CDS: Mobile element protein
contig_5	1768	1768	SNV	G	C/G	58.0/42.0	169	CDS: Mobile element protein
contig_5	1777	1777	SNV	A	C/A	56.7/43.3	164	CDS: Mobile element protein
contig_5	1796	1796	SNV	T	T/G	69.4/30.6	157	CDS: Mobile element protein
contig_5	1798	1798	MNV	AC	AC/CA	69.0/29.9	281	CDS: Mobile element protein
contig_5	32344	32344	InDel	-	CGAACGGTTTTTGTTCGTT	92.3	26	
contig_5	47690	47709	SNV	G	A	100	64	CDS: Mobile element protein
contig_6	170837	170837	InDel	-	GCCAAACTTAGGCAT	100	47	
contig_6	475043	475058	SNV	C	T	100	37	CDS: ABC transporter EcsA
contig_6	475145	475160	InDel	T	-	96.9	32	CDS: ABC transporter EcsA
contig_1	957	957	SNV	T	T/G	82.4/17.6	17	CDS: hypothetical protein
contig_1	968	968	MNV	AAG	AAG/TGC	61.3/25.8	31	CDS: hypothetical protein
contig_1	977	977	SNV	C	C/G	73.7/26.3	19	CDS: hypothetical protein
contig_1	981	981	SNV	C	C/T	60.0/40.0	20	CDS: hypothetical protein
contig_1	986	986	SNV	A	A/T	60.0/40.0	20	CDS: hypothetical protein
contig_1	988	988	SNV	T	T/C	60.0/40.0	20	CDS: hypothetical protein
contig_1	992	992	MNV	CT	CT/TC	59.0/41.0	39	CDS: hypothetical protein
contig_1	995	995	SNV	A	A/C	61.9/38.1	21	CDS: hypothetical protein
contig_1	1000	1000	MNV	TT	TT/GC	55.1/44.9	49	CDS: hypothetical protein
contig_1	1004	1004	SNV	A	A/G	56.0/44.0	25	CDS: hypothetical protein
contig_1	1007	1007	SNV	T	T/C	52.0/44.0	25	CDS: hypothetical protein
contig_1	1010	1010	SNV	G	G/A	54.2/45.8	24	CDS: hypothetical protein
contig_1	1022	1022	SNV	G	G/A	57.7/42.3	26	CDS: hypothetical protein
contig_1	1025	1025	SNV	T	T/C	57.7/42.3	26	CDS: hypothetical protein
contig_1	1028	1028	MNV	TTC	TTC/GCT	55.1/44.9	49	CDS: hypothetical protein
contig_1	1070	1070	InDel	-	-/GGTTG	70.2/29.8	47	CDS: hypothetical protein
contig_1	1072	1072	MNV	CA	CA/TTC	70.2/25.5	47	CDS: hypothetical protein
contig_1	1097	1097	SNV	C	C/T	69.6/30.4	23	CDS: hypothetical protein
contig_1	1103	1103	SNV	A	A/C	69.6/30.4	23	CDS: hypothetical protein
contig_1	1115	1115	SNV	T	T/A	71.4/28.6	21	CDS: hypothetical protein
contig_1	1125	1125	SNV	C	C/G	77.3/22.7	22	CDS: hypothetical protein
contig_1	1132	1132	SNV	T	T/C	76.2/23.8	21	CDS: hypothetical protein
contig_1	11716	11716	SNV	T	T/G	70.7/29.3	92	CDS: T1SS secreted agglutinin RTX
contig_1	11718	11718	SNV	T	T/G	73.2/26.8	97	CDS: T1SS secreted agglutinin RTX
contig_1	11724	11724	SNV	C	C/A	72.8/27.2	92	CDS: T1SS secreted agglutinin RTX
contig_1	11728	11728	SNV	A	A/G	72.1/27.9	86	CDS: T1SS secreted agglutinin RTX
contig_1	11732	11732	SNV	G	G/C	74.2/25.8	89	CDS: T1SS secreted agglutinin RTX
contig_1	11737	11737	SNV	G	G/C	73.6/26.4	87	CDS: T1SS secreted agglutinin RTX
contig_1	11746	11746	InDel	C	C/-	73.5/26.5	83	CDS: T1SS secreted agglutinin RTX
contig_1	11752	11752	MNV	TT	TT/CG	74.5/22.7	141	CDS: T1SS secreted agglutinin RTX
contig_1	11755	11755	SNV	T	T/C	74.4/25.6	82	CDS: T1SS secreted agglutinin RTX
contig_1	182726	182726	InDel	-	ATT	100	91	
contig_8	72413	72413	SNV	A	T	100	58	
contig_9	3031	3031	SNV	C	C/T	78.1/21.9	311	rRNA: rRNA
contig_11	21	21	SNV	A	A/G	57.1/42.9	14	
contig_12	19	19	SNV	T	T/G	71.4/23.8	21	
contig_12	21	21	MNV	AC	AC/CA	73.7/26.3	38	
contig_12	7400	7400	SNV	C	T	100	58	
contig_12	7487	7487	SNV	C	T	100	53	
contig_14	1100	1100	SNV	A	A/G	56.3/43.8	112	
contig_14	1292	1292	SNV	C	C/T	55.6/44.4	117	

M8	contig_5	1231	1231	SNV	A	G/A	54.4/45.2	259	CDS: Mobile element protein
	contig_5	1314	1314	SNV	G	G/A	69.3/30.7	329	
	contig_5	1340	1340	SNV	G	T/G	61.7/38.3	316	
	contig_5	1363	1363	SNV	A	G/A	63.4/36.2	309	CDS: Mobile element protein
	contig_5	1387	1387	SNV	T	C/T	61.6/37.7	305	CDS: Mobile element protein
	contig_5	1389	1389	SNV	T	T/C	66.9/32.5	311	CDS: Mobile element protein
	contig_5	1434	1434	SNV	T	C/T	65.7/34.3	283	CDS: Mobile element protein
	contig_5	1442	1442	SNV	A	C/A	66.3/33.7	291	CDS: Mobile element protein
	contig_5	1486	1486	SNV	C	C/T	65.0/35.0	317	CDS: Mobile element protein
	contig_5	1534	1534	SNV	G	A/G	67.1/32.9	346	CDS: Mobile element protein
	contig_5	1537	1537	SNV	T	C/T	66.5/33.2	346	CDS: Mobile element protein
	contig_5	1716	1716	SNV	T	C/T	59.6/40.4	376	CDS: Mobile element protein
	contig_5	1768	1768	SNV	G	C/G	56.3/43.8	320	CDS: Mobile element protein
	contig_5	1777	1777	SNV	A	C/A	54.8/45.2	303	CDS: Mobile element protein
	contig_5	1796	1796	SNV	T	T/G	74.1/25.9	293	CDS: Mobile element protein
	contig_5	1798	1798	MNV	AC	AC/CA/AA	72.4/27.0/0.4	496	CDS: Mobile element protein
	contig_5	19081	19081	SNV	C	A	100	75	CDS: TraG/TraD family
	contig_5	32344	32344	InDel	-	CGAACGGTTTTTGTTCGTT	94.5	73	
	contig_5	47690	47709	SNV	G	A	100	101	CDS: Mobile element protein
	contig_6	170837	170837	InDel	-	GCCAAACTTAGGCAT	90.9	99	
	contig_6	474315	474330	MNV	GCC	T	97.1	210	CDS: ABC EscB
	contig_1	1346	1346	InDel	-	-/A	76.9/21.2	52	CDS: hypothetical protein
	contig_1	1352	1352	InDel	G	G/-	82.0/18.0	50	CDS: hypothetical protein
	contig_1	1354	1354	SNV	T	T/C	81.3/18.8	48	CDS: hypothetical protein
	contig_1	1359	1359	MNV	TCG	TCG/CTA	70.4/25.4	71	CDS: hypothetical protein
	contig_1	1364	1364	SNV	G	G/A	80.9/19.1	47	CDS: hypothetical protein
	contig_1	1367	1367	SNV	A	A/G	79.2/18.8	48	CDS: hypothetical protein
	contig_1	1373	1373	SNV	T	T/C	79.2/18.8	48	CDS: hypothetical protein
	contig_1	1379	1379	SNV	C	C/A	79.6/18.4	49	CDS: hypothetical protein
	contig_1	1397	1397	SNV	C	C/A	78.7/21.3	47	CDS: hypothetical protein
	contig_1	11716	11716	SNV	T	T/G	73.0/23.0	152	CDS: T1SS secreted agglutinin RTX
	contig_1	11718	11718	SNV	T	T/G	75.8/24.2	153	CDS: T1SS secreted agglutinin RTX
	contig_1	182726	182726	InDel	-	ATT	99.3	153	
	contig_8	72413	72413	SNV	A	T	100	78	
	contig_11	21	21	SNV	A	A/G	69.7/30.3	33	
	contig_12	19	19	SNV	T	T/G	75.0/25.0	28	
	contig_12	21	21	MNV	AC	AC/CA	71.2/26.9	52	
	contig_12	7400	7400	SNV	C	T	100	113	
	contig_12	7487	7487	SNV	C	T	99.2	126	
	contig_14	1100	1100	SNV	A	G/A	51.4/48.2	247	
	contig_14	1292	1292	SNV	C	C/T	54.3/45.7	197	



Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no