A master thesis in microbiology for the Norwegian degree master of pharmacy

Events occurring during introgression of an antibiotic resistance gene in mutator populations of *Acinetobacter baylyi*

Maryam Fazelpour
Supervisors: prof. Kaare M. Nielsen, Dr. Odd Gunnar Wikmark, Dr. Morten Andresen

Department of Pharmacy, Faculty of Medicine, University of Tromsø
Spring 2007

Acknowledgments

The work presented in this thesis was carried out in the period of February 2007 until November 2007 in the department of pharmacy at the section of pharmacology, university of Tromsø.

I am grateful to Prof. Kaare M. Nielsen my supervisor for allowing me to experience some aspects of the interesting world of research and for his support and guidance during my work.

Thanks to my supervisor Post.doc Odd G. Wikmark for his guidance.

I am grateful to Post.doc Morten Andresen for his encouragement, comments and guidance and for helping me to work structurally, especially during writing the thesis.

I would like to thank Ph.D. student Hege Sletvold for being such a wonderful person and lab partner, for always helping me with my questions, sharing her knowledge, routine and for her support. Thanks to all other members of the microbiology lab and especially to Ane L.G. Utnes for her support and friendship. Thanks to Anne H. Conradi for her patience and for helping me to obtain material and equipment.

To my wonderful family thanks for always supporting and encouraging me during hard times, for your confidence and unconditional love.

I am thankful to my daughter Jasmin for her patience and understanding. You are the biggest happiness in my life and I am proud to be your mum.

Abstract

Natural transformation is recognized as a major horizontal gene transfer mechanism which impacts the genetic adaptation, diversity and evolution of prokaryotes. Natural transformation is the uptake of extracellular DNA and its integration in a host genome. The fact that species take up foreign nucleotide sequences is believed to be followed by elimination of non-advantageous DNA sequence since bacterial genomes do not grow in size. Transfer of a gene from one species into the gene pool of another by backcrossing of a hybrid transformant with one of the parents is defined as introgression. In this study the hypothesis that bacteria during introgression might regulate acquisition of foreign DNA by eliminating the non-selected sequence by recombination was examined. I specially wanted to expand on previous studies by using DNA sequencing technologies and the *Acinetobacter baylyi* BD413 (ADP1) published genome sequence to examine events that occur at DNA level during introgression in hybrid transformants. *A. baylyi* strain ADP1 is a highly recombinogenic bacterium and known to be competent for natural transformation. This thesis contributes to earlier (Ray, 2007) and ongoing projects investigating the fate of non-selected DNA during introgression of a chromosomal antibiotic resistance gene (*npII*) in the bacterium *A. baylyi*. Hybrid transformants of *A. baylyi* ADP7021 contained a kanamycin resistant gene, a *mutS*-gene deletion and an undetermined amount of the donor strain, *A.* sp. 62A1. The mutator strain *A. baylyi* ADP 7021 (Δ*mutS*) was used as recipient to determine how methyl-directed mismatch repair affects the process of introgression in this species. The generation of the initial heterogamic transformation and ten subsequent back cross transformations were performed in a previous study (Ray, 2007). DNA flanking the nptII insertion from the first and the 10^{th} generations was sequenced and analyzed in this study. Sequencing of DNA by primer-walking and contig assembly and analysis by Sequencher 4.1.4 software were the main techniques used in this study. The aim was identifying recombination junctions and the insert size of the donor DNA. The results were compared to the complete genomic sequence of ADP1 using the software BLAST. Putative recombination junctions were identified in all ten back-cross transformants in this study and differences among the first and the 10^{th} generations of *A. baylyi* transformants were observed. It was clear that back-cross transformants harboured a significantly shorter donor sequence when compared to the initial generation. This indicates that during introgression, the non-selected sequence will be gradually eliminated. The effect of inactivation of mismatch repair system in the recombination process of transformants of this study comparing to the wild type could not

be determined because similar studies of the wild type recipient are still in process (Wikmark, unpublished).

Table of contents

Abbreviations

BLAST: Basic Local Alignment Search Tool

***nptII*-gene**: Neomycin phosphotransferase II-gene

KANR: Kanamycin resistance

kb: Kilobasepairs

bp: basepair

LB: Luria Broth Base

LB-SSK: Luria Broth medium, containing antibiotics Streptomycin and Spectinomycin

HGT: Horizontal gene transfer

DNA: Deoxyribonucleic acid

Exo-SAP-IT: Exonuclease I - Shrimp alkaline phosphatise

Exo.AP: Exonuclease I-Antartic phosphatise

EBI-EMBL: European bioinformatic institute-European molecular biology laboratory

MMR: Methyl-directed mismatch repair

MutS: Protein that recognize and binds to mismatches during DNA replication

Δ*Mut*S: Deletion in *mutS*-gene

dNTPs: Deoxyribonuclotides triphosphates (dATP, dGTP, dCTP and dTTP)

NCBI: National Center for Biotechnology Information

PCR: Polymerase chain reaction

TAE: Tris-acetat-EDTA

TBE: Tris-borat-EDTA

Introduction

Horizontal gene transfer (HGT)

Horizontal gene transfer has been defined as the movement of genetic material between phylogenetically unrelated organisms under specific conditions (Thomas and Nielsen, 2005). This happens by mechanisms other than parent to offspring and is heritable in which the information is transferred through progeny as the cell divides. Horizontal gene transfer has an important role in enabling an organism to acquire new capabilities such as resistance to antibiotics. Transfer of genetic material between bacteria/microbes will impact the development of new traits in bacteria. It has been suggested that "any biological advantage provided to the recipient organism by the transferred DNA creates selective pressure for its retention in the host genome" (Garcia-Vallve *et al*., 2003). Horizontally acquired chromosomal DNA or mobile genetic elements such as plasmids and transposons that provide functional and metabolic benefits to the host are capable to spread rapidly within bacterial population (Thomas and Nielsen, 2005). Some horizontal acquisitions might be effectively neutral, and their survival will depend on chance events. Successful gene transfer and insertion of foreign DNA into a recipient must be stable in the host genome to ensure transmission of the new sequence over generations. Integration of foreign chromosomal DNA into a recipient bacterial chromosome is depended on many factors and barriers, some factors are mentioned below.

Genetic transformation

DNA can be transferred between bacteria by at least three different mechanisms: transformation, transduction and conjugation. Transformation is the uptake of free DNA from the extracellular medium or from the environment by prokaryotes and eukaryotes. It can occur naturally within competent microorganisms or may be performed artificially in laboratory by exposure to specific chemicals. In bacteria, the process begins when a bacteria cell (living or dead) releases DNA into the surrounding medium. Extracellular, naked DNA is vulnerable to degradation in the environment by factors such as temperature and pH-variation but it may encounter another bacterial cell before degradation. If the foreign DNA is recognized by the host cell, the cell may take transport it across the cell wall and cell membrane and allow it to recombine with its gene material. The resulting recombinant cell is called a ***transformant.***

Natural transformation is a mechanism of gene transfer that allows uptake of genetic material from diverged species in bacteria (Lorenz and Wackernagel 1994, Nielsen *et al.,* 2000). It occurs under natural condition which is normally compatible to the bacterial natural growth environment and can occur between distantly related organisms. Some species, for example *Acinetobacter baylyi* can develop a characteristic termed as "competence" to take up DNA from related or non related organisms. This characteristic is not a common feature and differs within different organisms. Incorporation of an acquired sequence into the genome of a recipient is depended on passing a number of cellular and extra cellular barriers. Prerequisites for natural transformation among other factors are, release of naked DNA, binding of DNA from the donor to the cell surface of the recipient cell, transported into the bacterial cytoplasm, passing through (DNA from the donor) restriction and anti restriction barriers of recipient, integration and stabilization of the foreign DNA into the host genome. Some bacterial species take up DNA from both close related strains and completely unrelated bacteria. Examples include *Bacillus subtilis*, *Azotobacter vinelandii* whereas some bacteria take up DNA only from closely related species such as *Neisseria gonorrhoeae* and *Haemophilus influenzae* preferentially take up DNA of their own species (Thomas and Nielsen 2005, De Vries and Wackernagel 2002, Salyers, 1995). Both plasmids and chromosomal DNA can be transferred into the recipient cytoplasm. Chromosomal DNA can integrate into the recipient's genome via homologous recombination in which the incoming DNA must contain similar regions to the host genome. In addition to homology-based recombination processes, natural transformation can also lead to illegitimate recombination events when there is little or no similarity between the recipient and the donor DNA, this is discussed below.

Recombination

Exchange of DNA sequences and transfer of genetic information occurs through an intricately regulated series of enzymatic reactions. DNA sequences may recombine with one another by mechanisms involving pairing of nucleotide sequences in DNA molecules and presences in which phosphodiester bond breaks and rejoins. This type of sequence rearranging and replacing is known as *genetic recombination*. Recombination in *E. coli* and other bacteria is mediated by *Rec*A which its activity is ATP (Adenosine triphosphate) depended and requires a minimal length of DNA homology between the donor sequence and a double strand recipient DNA. Homologous recombination requires sequence similarity between the incoming donor DNA and DNA sequence of the recipient. In *E. coli* a minimum of approximately 20 base pairs is acquired for successful recombination with circular DNA (Watt *et al,* 1985). In most cases,

DNA is integrated into the recipient genome through homologous recombination. Bacterial recombination can also introduce donor DNA that has a limited homologous counterpart in the recipient. Hence it is termed heterologous recombination. A heterologus donor segment that is flanked by DNA that is homologous with the recipient can integrate into the recipient genome via homologous recombination, occurring in the flanking regions (double-crossover) (Harris-Warrick and Lederberg, 1978; Hamilton *et al*., 1989; Majewski and Cohan, 1999).

Illegitimate recombination

Heterologous recombination events requiring between 150-200 bp DNA similarity, have been named homology-facilitated illegitimate recombination (De Vries and Wackernagel, 2002). Illegitimate (nonhomologous) recombination events are associated with very small sequences ("anchors") that join DNA molecules at sites with no or a few identical nucleotide sequences. These short sequences can be as small as 3 or 8 base pair and do not require *Rec*A protein to be initiated. Such events are rare and happen in low frequency among bacteria (De Vries and Wackernagel, 2002).

Methyl directed DNA mismatch repair

During the processes of replication, DNA mismatches between basepairs can occur. When gene transfer by recombination of an incoming exogenous DNA from environment into a recipient organism goes through successfully, it results in integration of the new foreign sequence in the host genome. This new sequence must also, pass the mismatch repair system. Correction of mismatched nucleotide pairs in the recipient genome take place by activity of a range of proteins and enzymes. The repair pathway is encoded by *mutH, mutS, mutL* and *mutU* genes in *E. coli* among other mechanisms of repair. These genes are necessary for editing bases during replication and recombination and it is believed that these genes function as a complex. The repair system in the process of DNA replication initiates when the mutS protein recognizes a mismatch base, mutL is thought to act as a linker and binds mutH to mutS. The mutH protein binds at a hemimethylated (GATC) sequence and causes a nick to cleave the non-methylated strand. These genes were originally identified as mutator genes since mutations in these genes will result in defective repair systems with the consequence that the cells will have a higher rate of spontaneous mutation than usual. MutHLS methyl directed mismatch repair edit the "wrong", newly synthesized or recombined base because it assumes that the sequence of the parental template strand is correct and the daughter strand must be repaired. Thus the mismatch repair system is considered as an inhibitor of recombination between non-identical sequences

and reduces mutation rate in bacteria. Mismatch repair enzymes distinguish the newly inserted sequence (daughter strand) from the parent strand because the new sequence is not yet methylated by modification enzymes (in *E. coli*; DNA adenine methylase). In this way, MMR proteins resolve the irregularity in DNA caused by a mismatch between the parental strand and the daughter strand and the gap is replaced using the parental strand as a template for synthesis. In some species such as *E. coli* and *Salmonella enterica*, mismatch repair acts as a powerful force of sexual isolation between these species where inactivation of MMR genes increases the frequency of mutations as well as interspecies recombination. Thus, deletion of the MMR genes including the *mutS*-gene can enable diverged bacteria to undergo horizontal gene transfer and genetic exchange (Matic *et al*., 1996). In studies of *E. coli* and a mutator phenotype with *mutS*-deletion it was concluded that restoration of a functional *mutS*-gene by recombination was necessary for long term survival of the organisms (Denamur *et al.,* 2000, Brown *et al.,* 2001). In the present study, a mutator type of *A. baylyi* with *mutS*-deletion has been used to investigate events occurring during natural transformation and introgression of an antibiotic resistance gene.

Acinetobacter baylyi

Barbe *et al.* in 2004 described the characteristics of the genus *Acinetobacter*: "The Gram-negative bacteria which are now classified in the genus *Acinetobacter* can be distinguished by the following characteristics: they are oxidase-negative, catalase-positive, and strictly aerobic and possess a strict respiratory metabolism. They are immobile with no flagella, do not form spores and appear as cocci under the microscope in stationary phase or as short bacilli, often in pairs or assembled into chains". *A.* sp. strain ADP1 is closely related to representatives of *Pseudomonas aeruginosa* and *Pseudomonas putida*. *A.* sp. ADP1 has been a model organism for several studies of natural transformation and HGT because it is a highly competent strain and easy to transform. The encapsulated strain BD4 (stands for Butane-diol) was first isolated from soil by Taylor and Juni in 1961(Genoscope - Centre National de Séquençage http://www.genoscope.cns.fr/externe/English/corps_anglais.html). BD413 is a minicapsulated mutant of BD4 obtained by ultraviolet irradiation and appeared to be very easy to transform. It is described under the name BD413 and ADP1 in Europe and in USA, respectively.

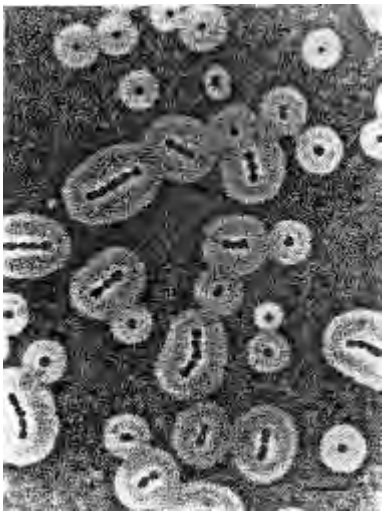


Figure 1. The figure shows *Acinetobacter* sp. BD4 cells which was isolated from soil. The figure is taken from the internet-site Genoscope (http://www.genoscope.cns.fr/externe/English/Projets/Projet_DY/organisme_DY.html) and is referred to studies of BD4 cells by Taylor and Juni, 1961.

Introgression

Introgression is movement and integration of genes from one organism to another by a process of backcrossing a hybrid with one of its parents. Introgression is a long term process and requires several generation of backcrosses to occur (Ray, 2007). The stable integration of the genes into another can occur within or between species. The introgressed sequence of bacteria will be adjusted to the base composition of host genome after some time. Chromosomal deletions act as a mechanism to eliminate genes that do not provide a meaningful function either a large or short acquired DNA. Hence gene acquisition is as frequent as gene loss.

Hybrid transformants used in this study

By using the *nptII*-gene (neomycin phosphotransferase II; kanamycin resistance) as selectable marker, in the previous studies (Ray *et al*., 2007) it was found that the integration of foreign DNA into the genome of the *A.* sp. BD413 during transformation was at last 10^{-9} fold lower than that of homologous DNA. Chromosomal DNA from *nptII*-gene insertion events (isolates) was used as donor DNA in filter transformations to determine the frequency at which these random marker gene insertions were transformed into recipient *A*. sp. BD413 (wild type) and *A. baylyi* ADP7021(mutator strain). For the present study, the mutator strain *A*. sp. ADP7021 was used as recipient and *A.* sp. 62A1 as donor strain. The bacteria transformant *A. baylyi* ADP7021 contains an undetermined amount of *A*. sp. 62A1 genomic DNA at the recombination locus.

The mutator strain *A. baylyi* ADP7021 (Δ*mutS*) transformant from this experiment was used as recipient to determine how methyl-directed mismatch repair affect the process of introgression in this species. To test the hypothesis that introgression is a mechanism by which non-selected transforming DNA is eliminated from recombining bacterial populations, the base sequence and fate of the invading DNA fragment inserted into the recipient was studied. The aim was to identify crossover junctions.

An overview of the bacteria strains construction and design is given in fig. 2. In this study, isolates from 0^{th} generation transformants (M16.4.0.1→5) and 10^{th} back-cross transformants (M16.4.10.1→10) of *A. baylyi* were individually cultured, DNA isolated and DNA flanking the nptII insertion was sequenced.

Construction of the strains used in this study

In this study, DNA sequences of *A. baylyi* ADP7021 transformants (Δ*mutS*) were analyzed. An overview of the construction (by Ray, 2007) of the strains for the introgression studies (by Ray, 2007) is given in figure 2.

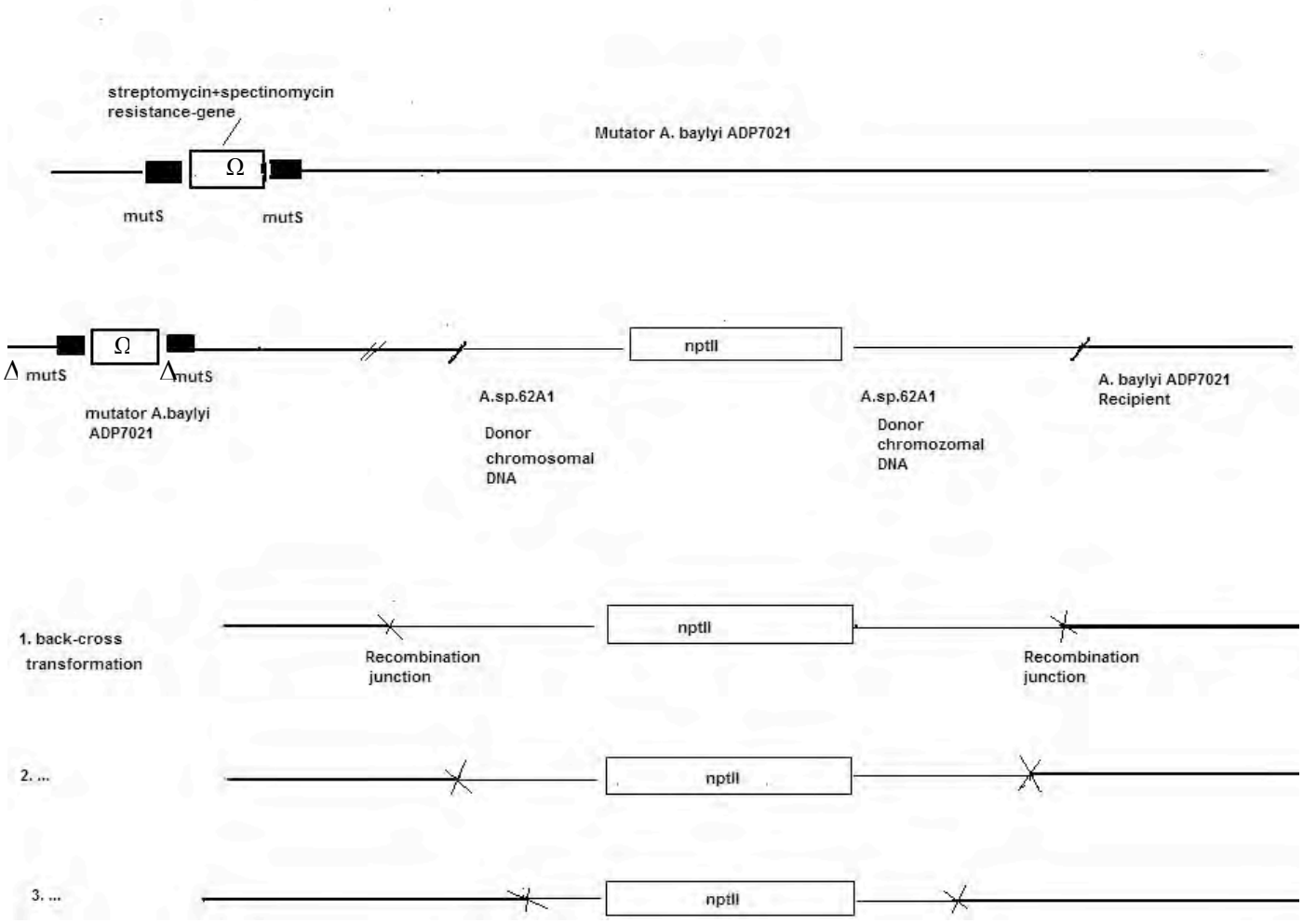


Figure 2. Schematic overview of mutator recipient strain *A. baylyi* ADP7021 transformants. The genome of the recipient strain has a transposome Ω-cassette: contains a *mutS* deletion in addition of the streptomycin and spectinomycin resistance gene (Young *et al*., 2001). The donor strain *A. baylyi* 62A1 DNA containing the kanamycin resistance gene (*nptII*) is transformed into the ADP7021 genome. Hypothetical flanking regions and cross-over junction in both sides of the *nptII*-gene are marked. The figure illustrates how repeatedly transforming (back-

crossing) hybrid transformants with the parents population result in that non-selected DNA from the donor strain may be eliminated from the genome.

Aims and objectives

Aim

To investigate the fate of non-selected DNA during introgression of a chromosomal antibiotic resistance gene (*nptII*) in the bacterium *Acinetobacter baylyi*. In particular to focus on the molecular events occurring during introgression in the mutator strain *A. baylyi* ADP7021.

Objective

To examine the sequences flanking the *nptII* gene in the initial transformants (0^{th} generation) as well as in the back-transformants from the 10^{th} backcross generation to identify putative recombination junctions on the right and left ends of the *nptII* gene in 62A1/ADP7021 transformants.

Materials and Methods

Natural transformation of Acinetobacter baylyi

The *A. baylyi* transformants used in this study were generated (Ray, 2007) in a previous study of natural transformation of *A. baylyi* with heterologous DNA and of introgression studies of the nptII-gene insertion in *A. baylyi*. An overview and illustration of the experimental approach is given in figure 3. The bacterial isolates that are used in the present experiment are *A. baylyi* 62A1/7021 transformants described as follows:

Chromosomal DNA from strain *A*. sp. 62A1 with an *nptII*-gene insertion was used to transform the mutator recipient *A. baylyi* ADP7021 by filter transformation. Up to 10 single transformant colonies was picked from these isolates. These are the initially transformed isolates, denoted initial transformants or 0^{th} generation: M16.4.0.1-10.

The letter M indicates the Mutator recipient *A. baylyi* ADP7021 and 16.4 refers to the donor isolate of *Acinetobacter* with the *nptII*-gene insertion (*A.* sp. 62A1). The number 1 refers to the first isolate from 0^{th} generation (and so on). Bacteria isolates from 0^{th} generation examined in this study are M16.4.0.1 – M16.4.0.5. Back crossing hybrid transformants was performed and all hybrid transformants colonies were collected from the initial single plate and combined into a cell mixture. Mixed hybrid transformants were cultured with selection and DNA was isolated. The chromosomal DNA was isolated and transformed into the recipient strain *A. baylyi* ADP 7021. From these transformation events, 10 colonies was picked and stored in medium (25% v/v glycerol in Luria Broth). These isolate were denoted back cross transformants and the process was repeated for ten generation. Back-cross transformant from the 10^{th} generation used in the present study are: M16.4.10.1-10. The number 10 refers to the 10^{th} generation. The number 1 refers to the first isolate from a collection of 10 back cross transformants. In summary, the transformants in this study are isolates from 5 initial, (0^{th} generation: M16.4.0.1→5) and 10 back-cross transformants, (10^{th} generation: M16.4.10.1→10). These transformants were individually cultured and DNA was isolated. DNA from these tranformants was isolated to be used for sequencing of DNA flanking the *nptII* insertion.

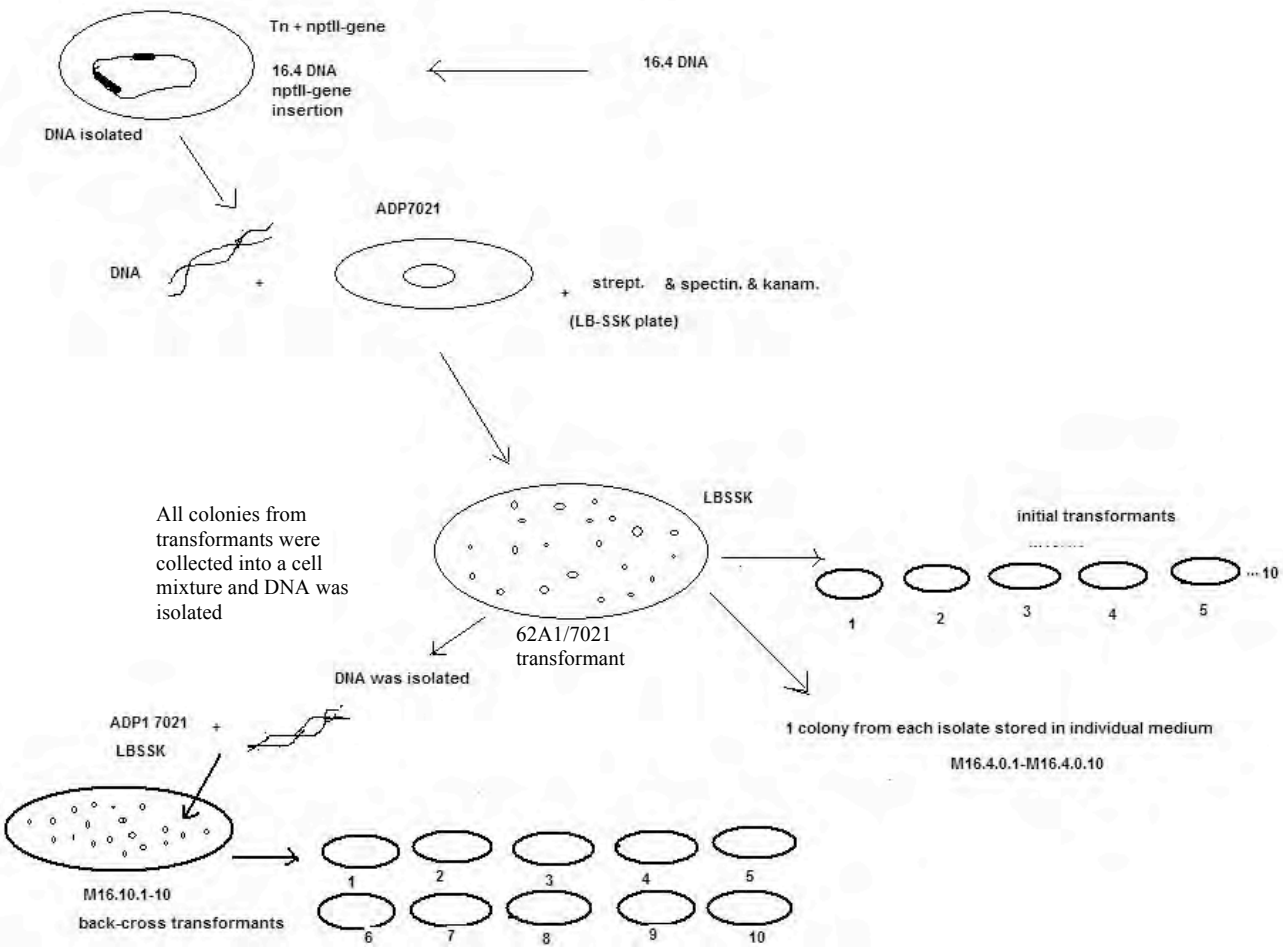


Figure 3. The *nptII*-gene was inserted into the chromosome of the donor strain 62A1 containing a *mutS* deletion and a transposon, which codes for streptomycin and spectinomycin resistance. DNA was isolated from each isolate. Transformation of chromosomal insertion events into the mutator recipient ADP7021 was the next step. Cells were collected and plated on medium. Selection was performed on LB-SSK (see below). One colony from each isolate called initial or 0^{th} transformants stored in individual LB-SSK medium. Hybrid transformants colonies were collected from the initial plate and combined into a cell mixture. DNA was isolated from the mixture and then used to transform the recipient strain. From these transformation events 10 colonies was picked, for ten back cross generations.

Culturing A. baylyi strain ADP7021

Growth media

The bacteria were grown on Luria Bertani plates with the antibiotics needed for selection at 32 ºC. LB-Agar was prepared by adding 13 g Agar (Merck, Germany) to 25 g Luria Broth Base (Invitrogen, Germany) and 1 l distilled water followed by autoclaving at 121 ºC for 20 min. Luria Broth (LB) was prepared by adding 25 g in 1 L distilled water. The media were autoclaved at 121 ºC for 20 min. Both media were added 1 ml kanamycin (Fluka sigma Aldrich, Germany) (50 mg/ml), 1 ml streptomycin (10 mg/ml) and 1.46 ml (spectinomycin 35 mg/ml) at a temperature of approximately 50 ºC (the medium will be denoted LB-SSK from now on). From the incubated overnight cultures on the LB-SSK-plates, one colony was picked up aseptically and transferred to 10 ml sterile tubes containing 3 ml LB-SSK-medium. The bacteria cultures were incubated overnight at 33 ºC with shaking at 200 rpm (the cap of the vial was left slightly loose). The overnight cultures were transferred to new 50 ml tubes containing 27 ml fresh LB-SSK medium and was incubated at 33 ºC for 6-7 hours with shaking at 200 rpm. Cells were pelleted at 4000 rpm for 10 min at 4 ºC.

Cell lysates - PCR templates

This method was used to generate DNA from the transformants to perform PCR-analysis.
A freeze culture from each population (in 25% glycerol) stored at –70 ºC was thawed on ice for a few minutes and streaked on LB-SSK-agar plate. Plates were incubated at 33 ºC for 2 days. One colony from each isolate picked and transferred to a tube containing 100 µl of 0.9% NaCl and dissolved with a sterile inoculation needle. The suspension was boiled at 100 ºC for 10 min. followed by centrifugation at 13000 rpm for 5 min. The DNA containing supernatant was transferred to a new microfuge tube. The lysates were stored at –20 ºC. Samples of cell lysates were examined with variations and modifications in preparation.

Isolation of genomic DNA

The genomic DNA was isolated according to the Qiagen genomic DNA isolation handbook (Qiagen, Germany) with minor alterations as described below.

DNA isolation using the QIAGEN genomic tip 100/G column

Cultures were generated by diluting an overnight culture from one colony of bacteria in 3 ml fresh LB-SSK and were incubated at 33 ºC with 200 rpm shaking overnight. The overnight cultures were added 27 ml fresh LB-SSK medium in 50 ml tubes and incubated at 33 ºC with

200 rpm shaking for 6-7 hours. Cells were pelleted by centrifugation for at 4 ºC at 4000 rpm for 10 min. Preparation and composition of buffers and material is described in table 1. Pellets were resuspended in 3.5 ml B1 buffer and 7 µl of 100 mg/ml RNase A (Sigma, Germany) by vortexing at max speed for 10-15 min. Samples were added 80 µl of stock solution 100 mg/ml Lysozyme and 100 µl of stock solution 10 mg/ml Proteinase K (Sigma, Germany). The samples were incubated in a water bath at 37 ºC for 60-90 min. After observing if samples were transparent, they were added buffer B2 followed by vortexing and then incubation at 50 ºC for 30-40 min. Buffers B1, B2, QBT, QC, QF were pre-warmed to 50 ºC before use. Qiagen 100G genomic tips were equilibrated with 4 ml QBT buffer. Samples were vortexed for 10 seceond before loaded on the 100G genomic columns. All liquids were allowed to drain through the column by gravity flow. The genomic tips were washed with 2 x 7.5 ml of buffer QC. To elute the DNA, the columns were placed in 50 ml Falcon tubes and buffer QF was added to drain by gravity flow through columns. DNA was precipitate as visible white mass with the elution buffer when isopropanol (3.5 ml) was added and tubes were 18-20 times inverted to assemble DNA. An inoculation needle was used to recover and transfer the DNA to a sterile new micro-tube. The DNA was washed in 70% ethanol (0.5 ml) for a few seconds and transferred to new sterile tubes. Tubes containing DNA were dried on the heating block at 37 ºC for 5 min. Pellets were solved in 100 µl Tris-Cl buffer (pH 8) and stored over night at 4 ºC. Tubes were centrifuged for 40 min at 13000 rpm to remove carbohydrates. Supernatant containing pure genomic DNA were transferred to new tubes. The concentration was measured and adjusted by adding Tris-Cl buffer (pH 8).

Table 1. Materials used for isolation of genomic DNA from bacterial cultures are listed below.

Component	**Composition**
Lysozyme	100 mg per ml in dH_2O
RNAse A	100 mg per ml in 1 x TE buffer pH 8.0
Proteinase K	10 mg per ml in dH_2O
B1 (bacterial lysis buffer)	50 mM Tris-Cl, pH 8.0 50 mM EDTA, 0.5% Tween®-20, 0.5% Triton X-100
B2 (bacterial lysis buffer)	3 M guanidine HCl, 20% Tween-20
QBT (Equilibration buffer)	750 mM NaCl, 50 mM MOPS pH 7.0, 15%

	Isopropanol, 0.15% Triton X-100
QC buffer (wash buffer)	1.0 M NaCl, 50 mM MOPS pH 7.0, 15% Isopropanol
QF buffer (Elution buffer)	1.25 M NaCl, 50 mM Tris-Cl pH 8.5, 15% isopropanol
TE (DNA storage)	10 mM Tris-Cl pH 8.0 1 mM EDTA pH 8.0
Ethanol	100 ml 70% v/v EtOH to dH_2O
Isopropanol	100% solution

DNA concentration

In all 15 *A. baylyi* transformants DNA isolations were checked for concentration and quality using the NanoDrop® ND-1000 Spectrophotometer. Isolated DNA was dissolved in 100 μl Tris-Cl buffer 10 mM (pH 8) and concentration was measured by absorbance at 260 nm. Concentrations of DNA in each tube were adjusted to ~ 350-600 ng/ul of DNA. The purity of DNA in the samples was suggested to be in an acceptable area with a ratio OD_{260} /OD_{280} of about 1.8-1.9. DNA-suspensions were stored at -20ºc.

Polymerase chain reaction (PCR)

In this study PCR was used to identify the flanking region of donor DNA fragment by using PCR products from both bacterial cell lysates and genomic DNA.
Temperature and cycling programs were optimized for each reaction, template and primer pair. A general cycling instruction for standard PCR program is given in table 2.

Table 2. General cycling instructions for a standard PCR program:

Temperature	**Time**	**Cycle**
94 ◦C	1-10 min	1
94 ◦C	15s-2 min	Repeat 30-40 times
40-65 ◦C	15-60 s	
72 ◦C	1-2 min	
70 ºC-74 ºC	5-15 min	1
4-15 ◦C	∞ (hold)	

PCR of DNA present in cell lysates

For the cell lysates several protocols with variation in concentration of template and primers were performed. Samples were set up by a mixture of 1 U DyNazyme II master mix (Finnzymes, Finland) in 20 µl reaction with forward and reverse primer, template and sterile

water. One example of a protocol for reaction set up is described in table 3. Two µl of the supernatant of each individual bacteria cell was used as template in PCR reactions with the cycling reactions. The supplier of all primers in this study was Operon. Primers were dissolved in a small volume of distilled, sterile water (according to the Operons' datasheet) to make a concentrated stock solution of 100 µM. Concentration of primers used in reactions were adjusted following Finnzymes instructions.

Table 3. Reaction mixture set up for PCR templates using bacteria cell lysates

Material	**Component**	**Amount**
Master Mix	2 x DynazymeII*	10 µl
Primer I	LP1046 (10 µM)	1 µl
Primer II	LPcRv1 (10 µlM)	1 µl
Template	Bacterial cell lysate	2 µl
Sterile water	ddH2O	6 µl

*The 2x optimized DyNazyme buffer contains DNA polymerase and 1.5 mM $MgCl_2$ and 200 µM dNTP in final reaction concentration.

PCR with genomic DNA template

Preparation of samples was performed with 1 µl of genomic DNA isolated by Qiagen 100/G column as described in previous section. Reaction mixtures included reverse and forward primers, respectively 16.4Rev-12 and LPcRfw4 (Operon), concentrations were modified according to DyNazyme II (Finnzymes, Finland) PCR master mix protocol. Samples with a final reaction mix of 50 µl were added 1 U DyNAnzym II. Table 4 shows material used in the reaction mixture for sequencing genomic DNA. The cycling program was optimized for this reaction, called MAR-1 cycling program, program is given in table 5.

Table 4. Material that were used for PCR of genomic DNA using primers 16.4Rev-12 and LpcRfw4 to identify flanking region of DNA

Material	**Component**	**Amount**
DyNazyme II DNA	DNA polymerase*	25 µl
Primer I	16.4Rev-12 (5 µmol)	1 µl
Primer II	LPcRfw4 (5 µmol)	1 µl
Template	Genomic DNA (~500 ng/µl)	1 µl
Sterile water	ddH2O	22 µl

*The 2 x optimized DyNazyme buffer supplies 1.5 mM $MgCl_2$ and 200 µM dNTP in final reaction concentration.

Table 5. Program (MAR-1) that was used for PCR of genomic DNA

Temperature	Time	Cycles
94 ◦C	3:40 min	1
94 ◦C	20 sec	Repeat 29 times
60 ◦C	30 sec	
72 ◦C	1:30 min	
72 ◦C	10 min	1
4 ◦C	∞ (hold)	

Further processing of PCR products from genomic DNA

Visualizing PCR product by gel electrophoresis was performed with the following concentration: onto the gel by running along with 5 µl ladder, 1 kb plus (Invitrogen, Germany).. After electrophoresis the gel was illuminated with an UV-trans-illuminator

Gel electrophoresis

Visualization of PCR products from both cell lysates and genomic DNA for was performed as follow: Samples were loaded on 0.7% agarose gel (Combrex Bio Science) added 0.5 µl/ml Ethidium Bromide (Sigma-Aldrich, Germany) in 1 x TBE buffer (Media laboratory, UiT). From the PCR-products 10 µl was added 2 µl of loading buffer 6 x T and loaded onto the gel along with 5 µl 1 kb plus ladder (Invitrogen, Germany). In addition a positive control (*A. baylyi* 16.4.5.5) was loaded parallel with the samples. The gel was run at 90 V for 55 min. After electrophoresis the gel was illuminated with an UV-trans-illuminator.

PCR product clean-up

PCR products generated from lysates were first treated with Exonuclease I and Antartic Phosphatase (Exo.Ap), according to the Exo-Sap-IT PCR clean-up protocol (USB, USA, modified by Hege Sletvold). Exo.Ap was used to clean up PCR products before sequencing. When PCR amplification is complete, unconsumed dNTPs and primers remaining in the PCR product mixture will interfere with the product. A master mix of Exo.Ap preparation is described in table 6. From this mixture, 1 µl was added to 5 µl PCR product. Primers and dNTPs were inactivated at 37 ºC in 30 min, and the enzymes were subsequently inactivated at 80 ºC in 15 min. Exo-Sap-IT program was run according to the cycling program in table 7.

Table 6. Material used for clean up PCR-product Exo-Sap-IT:

Component	**Composition**	**Amount**	**Supplier/references**
Antartic phosphatase	5000 U per ml	2 µl	New England Biolabs
NE Buffer for Antartic phosphatise	10 x concentration	1 µl	New England Biolabs
Exonuclease I	10 U per µl	1 µl	USB, USA
Sterile water	*dd*H2O	6 µl	IFA

Table 7. Program for Exo-Sap-IT

Temperature	**Time**
37◦C	30 min
80◦C	15 min
4 ◦C	∞

Cycle sequencing of cleaned up PCR products

Before sequencing, the final volume of 20 µl reaction mixtures were added forward or reverse primers (LP1046 or LPRv2) and ddH2O. For each sequencing reaction the following was mixed: 3.2 µl of primer, two different concentrations (3 and 5µl) of PCR-products, Big Dye 3.1 v (Applied Biosystems, USA*)* and $MgCl_2$ in individual 0.2 ml PCR tubes. Materials and concentration that were used for sequencing cleaned-up PCR-products generated (from lysates) are listed in table 8. Cycling program SEQ was run (table 9). Samples were sequenced by DNA-sequencing laboratory, Faculty of medicine, UiT.

Table 8. Reaction mixture for sequencing cleaned up PCR products generated from lysates.

Component	**Composition**	**Supplier**	**Amount**
Big Dye Terminator 3.1 version	*	Bio system, USA	3 µl
$MgCl_2$			4 µl
Primer (reverse/forward)	LPRv2 (10 µM) or Lp1046 (10 µM)	Operon	3.2 µl
Template	PCR-product (cleaned-up)		Variable

Sterile water	ddH2O	UiT	Variable

* Big Dye cycle Sequencing RR-100 composition: A-Dye, C-Dye, G-Dye, T- Dye – terminator; dATP, dCTP, dGTP, dUTP, Ampli Taq DNA polymerase FS with thermally stable pyrophosphate, $MgCl_2$ (5 μM)

Table 9. Program SEQ for sequencing PCR-product of cell lysate

Temperature	**Time**	**Cycles**
96 ◦C	3 min	1
96 ◦C	15 sec	Repeat 25 times
50 ◦C	10 sec	
60 ◦C	4 min	
4 ◦C	∞	1

Primer walk sequencing

After sequencing with the two known primers starting from the ends of *nptII*-gene nucleotide sequence, new DNA fragments were generated. The next step was designing primers derived from the ends of the newly determined DNA sequences. The process continued until the whole donor fragment was sequenced on the DNA of the transformants. New primers were designed both reverse and forward for each step of sequencing. Primers were ordered from Operon.

Primer design

Primer 3 (http://frodo.wimit.edu/), a software program was used for designing PCR primers. For the characteristics of primers in this study different parameters was considered for instance; avoiding complementary sequences within a primer sequence and avoiding 3 or more G or C at the 3'end. Lyophilized primers were dissolved in a small volume of distilled, sterile water (according to the Operons' datasheet) to make a concentrated stock solution of 100 μM. Primers were diluted to a concentration of 10 μM. All primers used and designed in this study are listed in table 14 and 15. Primers marked in the table that were used in preparing PCR products were chosen complementary to the genomic sequence of ADP1 while the rest of primers were designed from DNA fragments of the donor strain A. sp. 62A1. Primers Lp1046 and Rp152 were designed from nptII-gene nucleotide sequence in previous studies by Ray *et al.,* 2007.

Protocol for direct sequencing with concentration modification

Two different concentration modifications for reaction mixtures were used in this study. Reaction mixtures contained 4 μl primer (10 mmol/μl), 8 μl of Big Dye v. 3.1 (Biosystem, USA) and the isolated genomic DNA as template. The concentration of Big Dye and buffer 5x used in samples was adjusted according to the Biosystem instructions and were modified to supply a reduced volume of Big Dye by Morten Andresen later during the experiment. The modified reaction set up is outlined below in the table 10. Cycling program for direct sequencing was performed. Steps of the program are given in table 11.
Samples were sequenced in DNA-sequencing laboratory, Faculty of medicine, UiT.
Chromatogram of nucleotide sequences were visually inspected and edited by the computer program Sequencher version 4.1.4.

Table 10. Reaction mixture protocol for direct sequencing using genomic DNA

Material	**Component**	**Amount**
Big Dye Terminator 3.1 version	*	4 µl
Primer	Reverse or forward (10 µM) (listed in table 12 & 13)	2 µl
Buffer 5x	Sequencing buffer	2 µl
Template	Genomic DNA(350-500 ng/µl)	2 μl
Sterile water	ddH2O	2 μl

*Big Dye cycle Sequencing RR-100 composition: A-Dye, C-Dye, G-Dye, T-Dye-Terminator; dATP, dCTP, dGTP, dUTP, Ampli Taq DNA polymerase FS with thermally stable pyrophosphate, $MgCl_2$ (5 µM)

Table11. DIRSEKV, cycling program that was used for direct sequencing of genomic DNA

Temperature	**Time**	**Cycles**
95 ◦C	4 min	1
95 ◦C	30 sec	Repeat 99 times
55 ◦C	10 sec	
60 ◦C	1 min	
60 ºC	4 min	1
4 ◦C	∞ (hold)	

Nucleotide sequence analyzes

Assembly of DNA sequences was done using the computer program "Sequencher" (version 4.4.1). Sequences were assembled and compared, and inspections of differences or similarities

among transformants were analyzed. The program was used for building a contig of each individual DNA sequence and to analyze their chromatograms and bases. Assembly new sequences were followed by combining it into the contig to generate composition of the DNA sequences for all 15 *A. baylyi* transformants.

A consensus line (described in the next section) was created from these sequences to be used as a reference to detection differences in nucleotide positions and length of DNA insertions among transformants. At the completion of the project, DNA sequences were studied, analyzed, cleaned, and a contig map of all sequencing performed was edited for each transformant, see appendix 1. Each contig contains the *nptII*-gene fragment, the donor- and possibly, putative junction to the recipient DNA in both side of the gene

BLAST

This software is designed to take protein and nucleic acid sequences and compare them against a selection of NCBI databases (home page http://www.ncbi.nlm.nih.gov/BLAST/). DNA fragments of *A. baylyi* 7021/62A1 transformants were used as queries in BLAST. The query sequence was compared (aligned) with the database sequence of ADP1 and the amount of similarity was inspected.

Table 12. Primers that were used for sequencing *A. baylyi* transformants in the right side of the *nptII*-gene fragment (forward direction). Primers that were used in PCR products are included.

Forward primer (right side of nptII-gene)	Sequence 5'-> 3'	Tm ºC	Length	Position on the basis of *A.* sp. 62A1 sequence
ADPI23-S-I (PCR)*	CGTACACGAAAGGGCACAC	62.3	19	~2461
2LP1046 **	TTGAAGGATCAGATCACGCATCTTCCCG	~ 78	28	1046
16.4FW-1	ATGAATAAGGCCGGTTAGGC	60.4	20	717
16.4FW-2	TTCTGGTCACTAATTTAAAAAGTTCA	56.7	26	439
16.4FW-3	CTGCTATTGGTTTGGATAGTGG	60.8	22	1125
16.4FW-4	CGCGATTGTTCCAATTTCTT	56.3	20	1192
16.4FW-5	TGATTAAAAAGGGCGGGATT	56.3	20	1974
16.4FW-6	GATGCGACAAGAAACGAA	58.3	20	1912
16.4.FW-7	CGAATCATATAAGTCGCCAGA	58.7	21	2647
16.4FW-8	AACATGGATGCTGAGTTTGAA	58.7	21	2612
16.4.FW-9	GTGCTGAAAATGGTAATGCTG	58.6	21	3368
16.4FW-10	ATGCTACTGCAATGTATGCACC	60.8	22	3329
16.4.FW-11	GGATTGCCTACAAACCTCCA	60.4	20	3293
16.14.FW-12	ATCCACTCGTTGCACCTACC	62.4	20	3269
16.4.FW-13	CAGTGACCCTCTTGTTTGTAATG	61.1	24	3125

16.4FW14	GACCTATGCTACTGCAATGTATGC	62.8	24	3324
PFW-3**	ATAGTGGGGGATATCGTGGTGAGT	66.0	28	-
PF-4.1**	CGAAGTGGCAACTCCTGAAT	60.4	20	-
PF-4**	CTGGAGTTGATCAAGCCGCTACTG	66.2	24	-
16.4.FW-15	TGGAGTTGGGGTATTGGTTTAG	60.8	22	4158
16.4.FW-16	CCAGAGCTCCTCAAGAGTTGAT	62.6	22	3823
16.4.FW-17	AGTCAAATCCAGTTCCATCCTC	60.8	22	4980
16.4.FW-18	TGCAGTCTCAACACAGACAAA	58.6	21	4891
16.4.FW-19	TGGCTACTGATGACCCTCTT	62.6	22	3117
16.4.FW-20	AGACGGATGTGCAAGGTATTCT	60.8	22	3041
16.4.FW-21	GTTGACGCCTTTATCTTCTGCT	60.8	22	5650
16.4.FW-22	ATATTATTGGCACTCGCAGTT	58.9	22	5497
LPCRfw4*	GCGCCAAGATTGAAGGATTA	58.3	20	PCR
16.4.FW-23	ATATGGCTACGGCATTAACTGG	60.8	22	6289
16.4.fw-24	TGGTGATTTTACAGCCAACTCAA	58.9	22	6416

*Primer used in PCR product. The primer was designed on the basis of ADP1 sequence. The exact position on ADP1 genomic sequence was not obtained.
** The position and or Tm were not determined. The position of primer Lp1046 is given in the basis of *nptII*-gene fragment, from a previous study by Ray, 2007.

Table 13. Primers that were used for sequencing *A. baylyi* transformants in the left side of the *nptII*-gene fragment (reverse direction). Primers that were used in PCR products are included.

Reverse primer (left side of the *nptII*-gene)	Sequence 5'-> 3'	Tm ºC	Length	Position on the basis of *A.* sp. 62A1 sequence
LPCRrv1 (PCR)*	CTTCACCCAAATCCACCATC	60.4	20	~ 3000
LPCRrv2 (PCR)*	TACTCACGCCTGTCTGGTTG	62.45	20	~ 4000
2RP152 **	TTGAATAGGCTCATAACACCCC	~ 64	22	134
16.4.Rev-1	GTCTCCTCCCGATGATGATCCA	62.45	22	818
16.4.Rev-2	CCCAAGAATTTATTGCCGATT	56.71	21	1517
16.4.Rev-3	TGCTATTCAGCCAAGATGGTC	60.61	21	1419
16.4.Rev-4	CCGAAACTTCCAATCCTTGA	58.35	20	613
16.4.Rev-5	TGATGAACAGCACAGCAGAA	58.4	20	2225
16.4.Rev-6	CCAAACGCATATTCTGCTCA	58.4	20	2280
16.4.Rev-7	TACTTCGTCTGAGCCGTTACCC	64.54	22	3043
16.4.Rev-8	TAGCATCTGGGTCTGGATATAAGC	62.86	24	3123
16.4.Rev-9	TTTGACTAATACCAATGCACCAC	59.2	23	3760
16.4.Rev-10	AACATCCATCGCTAACCAATCTG	60.99	23	3928
16.4.Rev-11	CTTGACCAAATACAGCACGTTTAG	61.15	24	4014
16.4.Rev-12	ATCTAGGCATAGTTCCAAATGCTC	61.16	24	3694
16.4.Rev-13	TCTAGGGGCGTAACTTTTTGTC	60.81	22	4322
16.4.Rev-14	AACCACCAGGCACATAAATACC	60.81	22	4293
16.4.Rev-15	GGATGCTCCATGCCTACTTTAC	62.67	22	5113
16.4.Rev-16	CACATCGACTTGCTCACCTAAG	62.67	22	5030
16.4.Rev-17	ACAGCATCGTGACCTTCTACCT	62.67	22	5682

16.4.Rev-18	ACTAAAGCAGCCAGAACCAAAG	60.81	22	5595
16.4.Rev-19	CCAATTCATAAGGTGCAAACTG	58.94	22	6515
16.4.Rev-20	TAAATGCTGGTCCACAGGTCTT	60.81	22	6720

*Primer used in PCR product. Primers were designed in tha basis of ADP1 genomic sequence. The exact position on ADP1 genomic sequence was not obtained.
** Tm was not determined. The position of primer RP152 is given in the basis of *nptII*-gene fragment from a previous study by Ray, 2007.

Statistical analysis

The average inserted size of flanking DNA was calculated using Microsoft Excel t-test. A 95% confidence interval was found by testing the two hypotheses: H_0: The two means are equal and H_1: The two means are different. The probability p was chosen $p = 0.05$ (significance level). H_0 is accepted if $p > 0.05$.

Microhomology analysis

A microhomology comparison shows a graphical view of the similarity and differences between divergent multiple sequences and calculates the best mach of the sequences. The total achieved donor sequence in this study (from the consensus line) was compared with ADP1 genomic sequence (microhomology analyze) in a database, (http://www.ebi.ac.uk/Databases/nucleotide.html) EMBL- EBI- Clustalw, for alignment of the sequences. The result of such comparison between the donor, 62A1 from this study and the recipient ADP1, (The microhomology data-sheet) is not enclosed.

Results

Preparation of bacterial DNA

Two different methods for extracting DNA were examined. DNA to be used PCR from bacterial cell lysate and genomic DNA isolated by Qiagen 100/G Genomic-tips.

Experiments performed with DNA extracted from lysates

PCR product from bacterial DNA extracted from cell lysates were analyzed by gel electrophoresis. Migration of linear DNA through agarose gel was not observed. Expected product size was about 3000-4000 bp. To find possible source of errors, the method was modified by several variation as; changing PCR-program (temperature gradient), variation of concentration of template or primers and concentration of PCR products. In one case two bands on the lane by gel electrophoresis (fig. 4) were visualized. The two PCR products with band on lane 3 and 6 were cleaned up by Exo.Ap (according to Exo-Sap-IT protocol, USB, USA) for enzymatic removal of excess nucleotide and primers. Samples were sequenced to be identified but no sequence was obtained.

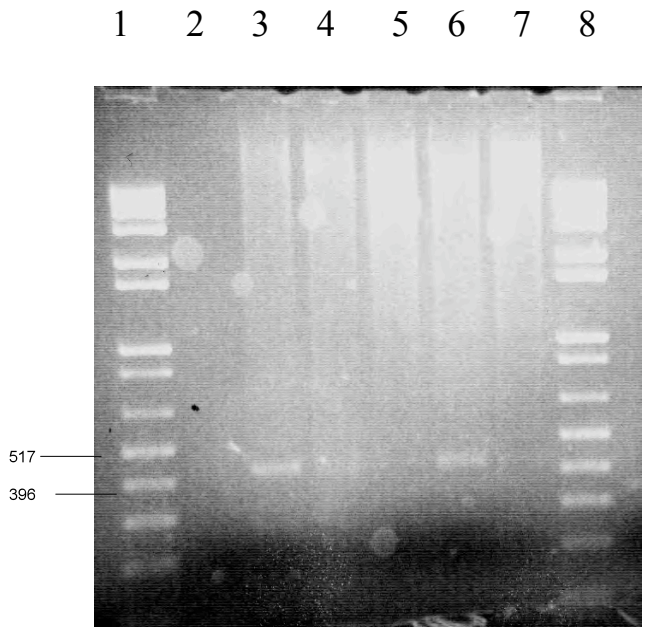


Figure 4. Agarose gel electrophoresis of PCR products from *A. baylyi* transformant prepared from cell lysates using primers Lp1046 and LPcRv1with expected product size about 3000-4000 bp. Lane 1 and 8: ladder 1 kb plus, lane 2: M16.4.10.1, lane 3: M16.4.10.2, lane 4:

M16.4.10.3, lane 5 M16.4.10.6. The result shows a band in lane 3 and lanes 6 with approximately size ~ 400-450 bp.

Direct sequencing- genomic DNA

Genomic DNA was isolated from 5 initial transformants and 10th generation back- cross transformants each by Qiagen genomic tip 100/G column. The concentration was measured by Nano-Drop and adjusted to ~ 350 - 500 ng/ul. Primer walking and direct sequencing of the genomic DNA was performed and samples were analyzed at the sequencing laboratory, Faculty of medicine, UiT. An overview of primers designed and used in this experiment is shown in table 12 and 13. Sequencing isolated DNA from transformants was started with the two primers (Lp1046 and Rp152) that their sequences overlapped the *nptII*-gene.

Sequence comparison

The ADP1 genome sequence has a small genome of 3.7 MB which has been completely sequenced. BLAST allowed comparing DNA fragments for transformants to the ADP1-sequence. Sequences with a similarity between 89-90% to the ADP1 genomic sequence were not accepted as the recipient gene. Primer walking was continued until 100 % match was achieved. Chromatograms (nucleotide sequences) generated were compared against each other and searched for alignments. An example from a chromatogram on Sequencher display is shown in 5 & 6. The figures show only a small part of the fragment of M16.4.10.3 chromatogram.

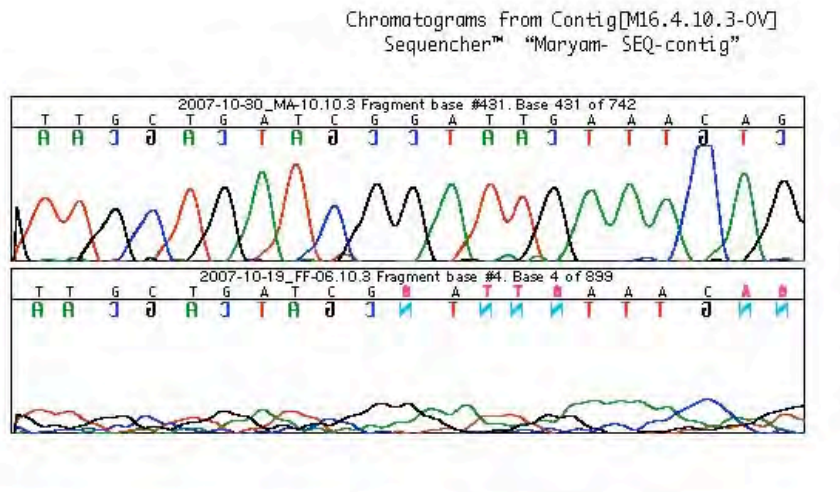


Figure 5. Chromatogram of upstream sequence (reverse) of the *nptII*-gene from *A. baylyi* M16.4.10.3 transformant contig at recombination junction inspected by Sequncher program version 4.1.1. The sequence in the bottom is identical with the sequence on the top but with lower clearness.

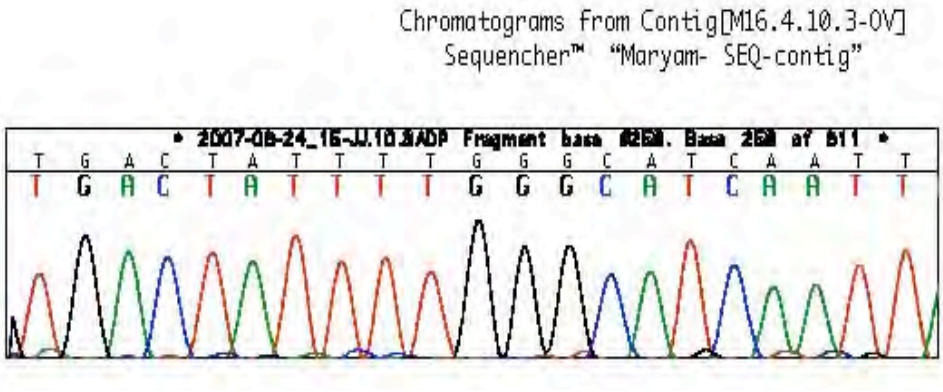


Figure 6. Chromatogram of downstream (forward) sequence of the *nptII*-gene from *A. baylyi* M16.4.10.3 tranformant contig at recombination junction. The chromatogram was inspected by Sequncher program version 4.1.1.

Consensus line of flanking regions

In bioinformatics, a consensus sequence represents the results of multiple sequence alignment for related sequences compared to each other. Consensus sequence shows which sequences are conserved and which sequences are variable.

A consensus line showing the calculated consensus of the sequences that were aligned was made in this study. After sorting the nucleotide sequences on both sides of *nptII*-gene, DNA fragments that did not show variability were assembled in a contig and were compared to the genomic sequence of ADP1. A reverse and forward consensus line was created which represented the DNA sequence of *A. baylyi* 62A1. Figure 7 & 8 shows the final consensus line obtained in this study and appendix 1 show all the consensus lines created in each step of the excessive primer walking done in this study. Primers designed on the basis of nucleotide sequences in this study are listed in table 12 -13. DNA sequences in right and left side of the *nptII* -gene and the putative flanking regions in 10^{th} generation was identified by primer walking. A contig consisting of the donor strain DNA sequence for each transformant in 10^{th} generation and 0^{th} generation was made. Appendix 1 shows Overlapping contig of all transformants as results from primer walk sequencing.

Primer-map

Primer map (figure 7 & 8) shows the schematic overview of positions of primer bases on the consensus sequence (nucleotide sequence of the donor strain) in this study. The forward and reverse primer-map containing a consensus line with all primer fragments was made. The primer-map demonstrates primers which generated a sequence clear enough to be inspected.

The starting primers Forward primer LP1046 and reverse RP152 overlapping the *nptII*-gene is not shown on the primer-map in figure 7 & 8.

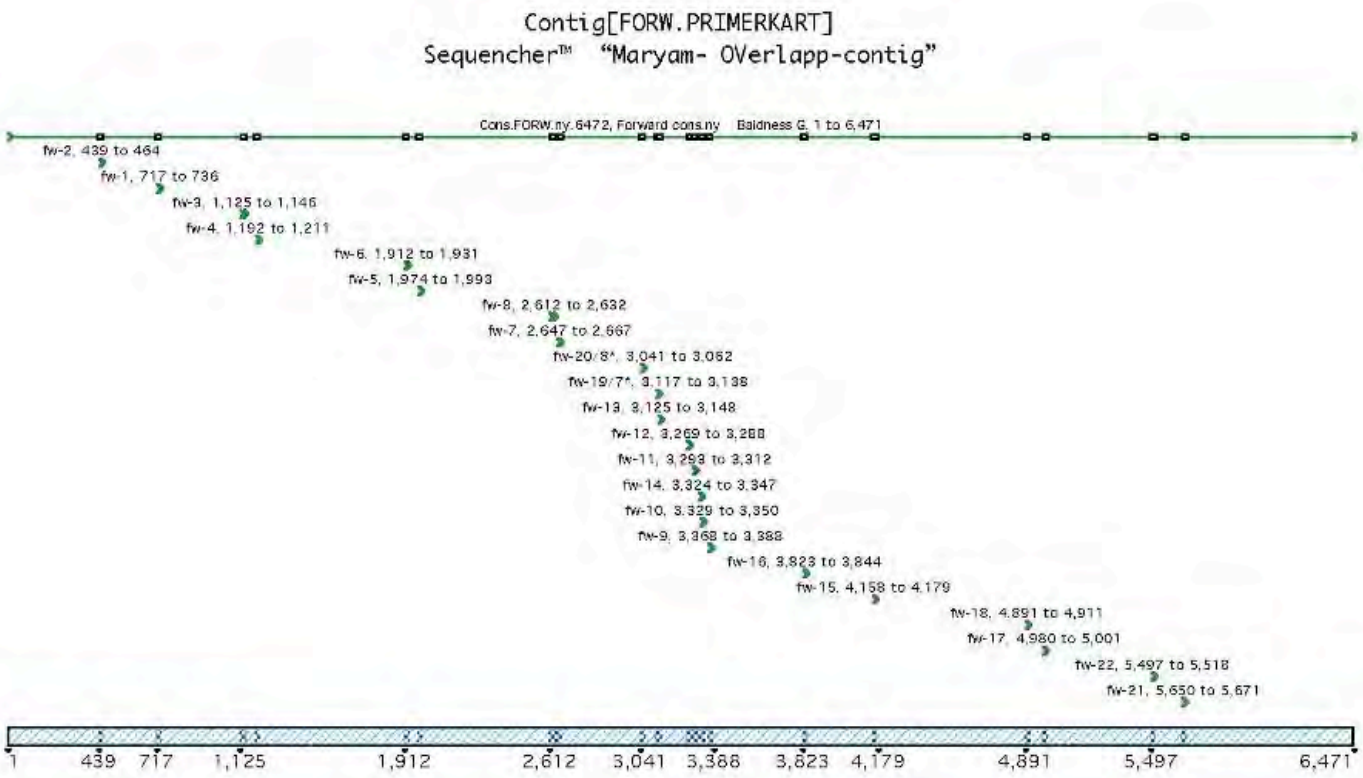


Figure 7. Forward primer-map. A consensus sequence (line) based on the length of *A. baylyi* 62A1 (donor) DNA fragment (7406 nucleotides) including the sequence produced from Lp1046 (133 bases overlap with *nptII*-gene) and all forward primers nucleotide sequence was created. The figure shows primers which have given a nucleotide sequence as the result of sequencing genomic DNA of 62A1/ADP7021 transformants in this study, starting with primer FW-2 ending with FW-24. The position of each primer fragment on the consensus line is shown in front of the name of primers.

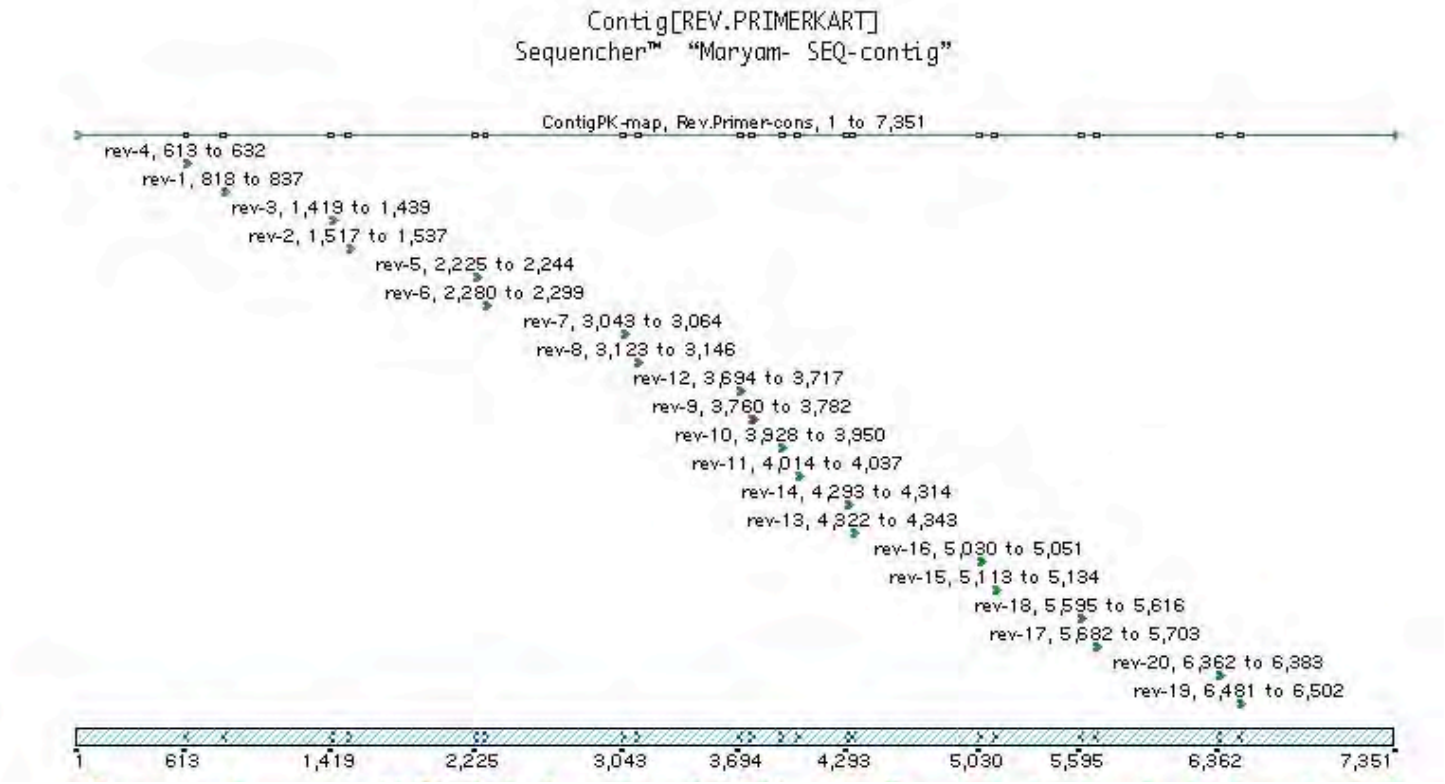


Figure 8. Reverse primer-map

A consensus sequence (line) based on the length of 62A1 (donor) DNA fragment (7351 nucleotides) and all reverse primers nucleotides sequence including primer RP152 (105 bases overlap with *nptII*-gene) was created. The figure shows primers which have given a nucleotide sequence as the result of sequencing the genomic DNA of 62A1/ADP7021 transformants in this study starting with primer rev-4 ending with rev-20. The position of each primer fragment on the consensus line is shown in front of the name of primers.

Identifying putative recombination junctions

Direct sequencing of genomic DNA from transformants and analyzing the results using Blast at the NCBI home page (http//www.ncbi.nlm.nih.gov/BLAST/) yielded a range of information. Sequences containing mismatched bases were alimented against the ADP1 genomic sequence on BLAST. Observation of "mismatches" was studied when they contained different bases in their sequence compared to sequences generated from the same primer among transformants. If alignment of the nucleotide fragment against ADP1 sequence scored 100%, it was assumed that the cross-over region was identified. When the first base mismatched observed the fragment was compared with the consensus line of the donor sequence and the position for the base was determined, as the cross-over junction.

Putative recombination junctions for all 10^{th} generation isolates were identified. Nucleotide positions of putative cross-over junctions are given in table 14 & 15. The specific flanking sequences for transformant cross-over junctions are shown in table 16 & 17.

Table 14. Nucleotide positions at left (reverse) and right side (forward) of the *nptII*-gene at cross over junctions in three initial transformant of *A. baylyi* ADP7021/62A1 and flanking sequence length of transformants.

Nucleotide position of cross-over junction (reverse)	**Flanking sequence length upstream of the nptII-gene**	**Reverse Primer**	**Transformant**	**Forward Primer**	**Flanking sequence length downstream of the nptII-gene**	**Nucleotide position of cross-over junction (forward)**
-	6880	16.4Rev-**20**	M16.4.**0.1**	16.4Fw-**2**	-	**977**
-	6900	16.4Rev-**20**	M16.4.**0.2**	16.4Fw-**24**	7480	-
5184	-	16.4Rev-**15**	M16.4.**0.3**	16.4Fw-**24**	7490	-
-	6839	16.4Rev-**20**	M16.4.**0.4**	16.4Fw-**24**	6850	-
-	6850	16.4Rev-**20**	M16.4.0.5	16.4Fw-**13 & 14***	-	**3394**

* Both primer 13 and 14 generated the same result. The table shows the last primer used for sequencing DNA among the non-identified junctions

Table 15. Nucleotide position at cross over junctions in 10th generation of *A. baylyi* ADP7021/62A1 transformants at left (reverse) and right side (forward) of the *nptII*-gene deteramined by primer walking.

Nucleotide position of cross-over junction (reverse)	**Reverse Primer**	**transformant**	**Forward Primer**	**Nucleotide position of cross-over junction (forward)**
1671	16.4Rev-**6**	M16.4.10.1	16Fw-16	**4298**
2401	16.4Rev-**6**	M16.4.10.2	16.4Fw-**5** &7,8	**2464**
5827	16.4Rev- **17**	M16.4.10.3	16.4Fw-5 ,7,**15**	**4464**
3141	16.4Rev-**12**	M16.4.10.4	Fw-5 &7 Fw**19***	**3502**
1705	16.4Rev-**2**	M16.4.10.5	16.4Fw-5& **7**	**3522**
1671	16.4Rev-**2**	M16.4.10.6	16.4Fw-**2**	**1091**
1671	16.4Rev-**2**	M16.4.10.7	16.4Fw-**2**	**1091**
1671	16.4Rev-**2**	M16.4.10.8	16.4Fw-**5**&7,8	**~2200**
1932	16.4Rev-**8** & (12)	M16.4.10.9	16.4Fw-**2**	**977**
2430	16.4Rev-**6,** 8,12	M16.4.10.-10	16.4Fw-**5** &7	**2465**

*Primer Fw19 was designed for the specific transformant M16.4.10.4 in the sequence region of 3000 bp.

The table shows the last primers used in sequencing reactions. The exact nucleotide position of the crossover junction of M16.4.10.8 was uncertain.

Nucleotide sequences at flanking regions

Table 16 & 17 shows the sequences at flanking regions and the first base mismatched (marked solid) in cross over position. The sequence on the top is the consensus line from the donor strain and sequence on the bottom is the cross over sequence belonging to the recipient`s genomic sequence.

Flanking regions are similar (identical sequences) at recombination junctions for transformants M16.4.10.6, and M16.4.10.7 in forward and reverse direction. Initial transformant M16.4.0.1 had an identical flanking sequence with isolate M16.4.10.9 from 10^{th} generation.

Isolate M16.4.10.8 has a uncertain sequence in flanking area and the exact position was not obtained. Sequences show that the same crossover junctions are appeared in different isolates. Nevertheless the sequences for integration appeared in different position of recipient genome among the majority of transformants.

Table 16. Nucleotide sequences in crossover regions for *A. baylyi* 7021/62A1 transformants on the left side of the *nptII*-gene, in reverse direction (R=reverse, con=consensus line made from *A.* sp. 62A1, seq=sequence of transformant). The sequence on the top is the consensus line from the donor strain and sequence on the bottom is the cross over sequence belonging to the recipient`s genomic sequence.

Transformant	Reverse DNA sequence (5→3)
M16.4.0.1	-
M16.4.0.2	-
M16.4.0.3	R. con. CCCTACTACGCATAAG**C**AAGTACGTATTTTGATTTTACGCG R. seq. TCCTACCACACATCAA**A**AAGTCCGTATTTTGATTTTACGCG
M16.4.0.4	-
M16.4.0.5	-
M16.4.10.1	R. con. CAACAATCATGACATT**C**TCTGCCTGATCTTTT R. seq. CGACAATCATGACATT**T**TCTGCCTGATCTTTT
M16.4.10.2	R. con. TTGCGGCCGCGGTAGC**C**TCTTTTGAAGATCAGGCT R. seq. TAGCTGCGGCAGTAGC**A**TCTTTTGAAGATCAGGCT
M16.4.10.3	R. con. ACATGGTTGTGAATATGC**T**GTTGTTGCTGATCGGATTG R. seq. ACATGGCTGTGAATATTC**A**GTTGTTGCTGATCGGATTG
M16.4.10.4	R. con. TTTTTATCCAGTTTATAG**T**CAGTTCTTTGGAAC R. seq. CTTTTATCCAGTTTATAG**C**CAGTTCTTTGGAAC
M16.4.10.5	R. con. TATTGGCATAAATGCTCA**A**ATCATTACGTAATAAATC R. seq. CTTCTGCATAGACACTTA**G**ATCATTACGTAATAAATC
M16.4.10.6	R. con. GTATTGGCATAAATGCTCA**A**ATCATTACGTAATAAAT R. seq. ACTTCTGCATAGACACTTA**G**ATCATTACGTAATAAAT
M16.4.10.7	R. con TATTGGCATAAATGCTCA**A**ATCATTACGTAATAAATCAAC R. seq. CTTCTGCATAGACACTTA**G**ATCATTACGTAATAAATCAAC
M16.4.10.8	R. con. AAATCAACAATCATGACATT**C**TCTGCCTGATCTTTTTGTGA R. seq. AGATCGACAATCATGACATT**T**TCTGCCTGATCTTTTTGTGA
M16.4.10.9	R. con. AATTTATTGAAACTTAATAC**G**AATGAAAACCCATATCCGCCA R. seq. AACTTACTTAAACTCAATAC**C**AATGAAAACCCATATCCGCCA
M16.4.10.10	R. con. TTGCAATTGCGGCCGCGGTAGC**C**TCTTTTGAAGATCAGGCTT R. seq. TTGCCATAGCTCCGGCAGTAGC**A**TCTTTTGAAGATCAGGCTT

Table 17. Nucleotide sequences in crossover regions for *A. baylyi* 7021/62A1 transformants on the right side of the *nptII*-gene, in forward direction (F=forward, con=consensus line made from *A.* sp. 62A1, seq= sequence of transformant). The sequence on the top is the consensus line from the donor strain and sequence on the bottom is the cross over sequence belonging to the recipient`s genomic sequence.

Transformant	Forward DNA sequence (5→3)
M16.4.0.1	F. con. CATATGACTATAGCTGTGG**T**TCAGCAGCGCTTACTACACT F. seq. CATATGACTATAGCTGTGG**A**TCTGCGGCATTAACAACATT
M16.4.0.2	-
M16.4.0.3	-
M16.4.0.4	-
M16.4.0.5	F. con. TGCTGAAAATGGTAATGCTGGCAA**C**TATATGGGACAGCAGG F. seq. TGCTGAAAATGGTAATGCTGGCAA**T**TATATGGGGCAACAAGT
M16.4.10.1	F. cons. GCGCAATTATTTGG**C**TTAGGTGTTGGATACCGA F. seq. GCGCAATTATTTGG**T**TTAGGTATTGGATATCGT
M16.4.10.2	F. con. CGAATTATGGATTTAAAGAAAC**G**GGTTTAGATCAAGCGCGTG F. seq. CGAATTATGGATTTAAAGAAAC**A**GGCTTAAATCAAAGACGTG
M16.4.10.3	F. con. AGATGTTAAAACCAATAC**C**AAAGTTACAGTACTCGG F. con. AGATGTTAAAACCAATAC**G**AAAGTGACTATTTTGGG
M16.4.10.4	F. con. GCTCTTGCTATGAAAAC**A**GATTTGCGATTTCCCAACGAGA F. seq. GCTCTTGCTATGAAAAC**T**GATTTAAGATTTCCCAATGAGT
M16.4.10.5	F. con. TTGCTATGAAAACAGATTT**G**CGATTTCCCAACGAG F. seq. TTGCTATGAAAACTGATTT**A**AGATTTCCCAATGAG
M16.4.10.6	F. con. CGGATTATTGAACGCCG**A**AGCTTCTCACTACTTGATATG F. seq. CGGATTATTGAACGCCG**T**AGTTTCTCACTACTCGATATG
M16.4.10.7	F. con. GCGGATTATTGAACGCCG**A**AGCTTCTCACTACTTGATATG F. seq. GCGGATTATTGAACGCCG**T**AGTTTCTCACTACTCGATATG
M16.4.10.8	F. con. GCTT**T**ACC:AGATAACTCATCTCAATTAAATCGTTTACGTA:TTG F. seq. GCTT**G**CCCGTGNNNNNNGATCCGAAATATGCCGAATTCNNNAC
M16.4.10.9	F. con. CATATGACTATAGCTGTGGTTCAGC**A**GCGCTTACTACAC F. seq. CATATGACTATAGCTGTGGATCTGC**G**GCATTAACAACAT
M16.4.10.10	F. con. AATTATGGATTTAAAGAAAC**G**GGTTTAGATCAAGCGCGT F. seq. AATTATGGATTTAAAGAAAC**A**GGCTTAAATCAAAGACGT

Length of flanking DNA in transformants

Initial transformants

In the 5 initial isolates (0^{th} generation), only 3 recombination junctions were identified; transformants M16.4.0.1 and M16.4.0.5 in forward direction and M16.4.0.3 in reverse direction. Therefore the length of the DNA sequence is equal to the total length in the specific direction. The length of DNA determined in these transformants is the total length of donor DNA (*A. baylyi* 62A1) acquired by the recipient, where as in non-identified junctions the flanking DNA length is continuing to be determined. Figure 9 shows length of donor DNA sequence for initial transformants on left and right side of the *nptII*-gene.

The length of the DNA sequences exceeds 7000 nucleotides in the right side (forward) of the *nptII*-gene. During primer walking some regions of the DNA sequence of the transformants were uncertain of this reason there are some gaps between DNA fragments in the contig assembly of some initial transformants (figures in appendix 1A). The gap between DNA fragments among some transformants can be observed around region of fw-3 primer in M16.4.0.1 contig. These gaps were the areas that fragments were not determined by primer walking and were complete by the consensus sequence.

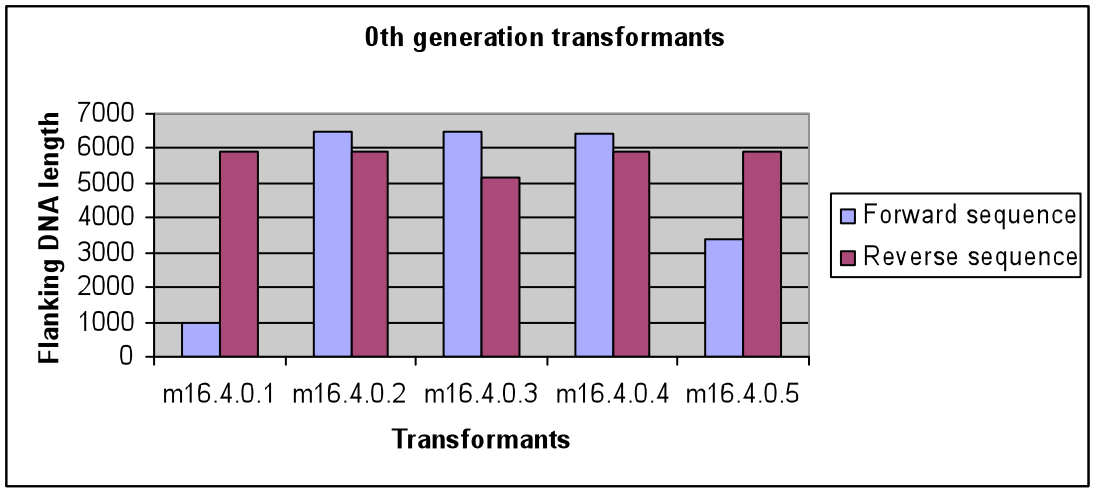


Figure 9. Transformants from *A. baylyi* 7021 were sequenced with different primers. The graph shows length of flanking DNA among initial transformants (0^{th} generation) on left and right side of the *nptII*-gene. The maximum length of M16.4.0.1 and M16.4.0.5 in forward- and M16.4.10.0.3 in reverse direction is equal to their recombination junction. All other transformants in this generation shows the length of donor sequence (*A. baylyi* 62A1) and recombination junctions are not reached.

Back - cross transformants (10^{th} generation)

The putative cross-over regions in 10 back-cross transformants were identified. Flanking DNA length in both reverse and forward direction was obtained. During primer walking some regions of the DNA sequence of transformants were uncertain of this reason transformant M16.4.10.8 the cross over region of the recipient DNA and the position of nucleotide is not certain. There are some gaps between DNA sequences among some other transformants. These gaps were the areas that fragments were not determined by primer walking and were complete by the consensus sequence. These sequences have been examined with different primers without generating fragments; examples are isolate M16.4.10.1 and M16.4.10.3 in which there is a gap in DNA fragment with respectively primer fw-7 to fw-16 and primer fw-3 to fw-5, see appendix1: figure 6&9 overlapping contig.

Figur 10 shows the length of DNA sequence for 10th back-cross transformants on the left and right side of the *nptII*-gene. Maximum DNA length is equal to the recombination junction. Only one transformant (M16.4.10.3 in reverse direction) has exceeded 5000 bases. The length of DNA determined in these transformants is equal to the total length of donor DNA (62A1) acquired by the recipient, DNA sequence acquired by the recipient determined in both side of *nptII*- gene.

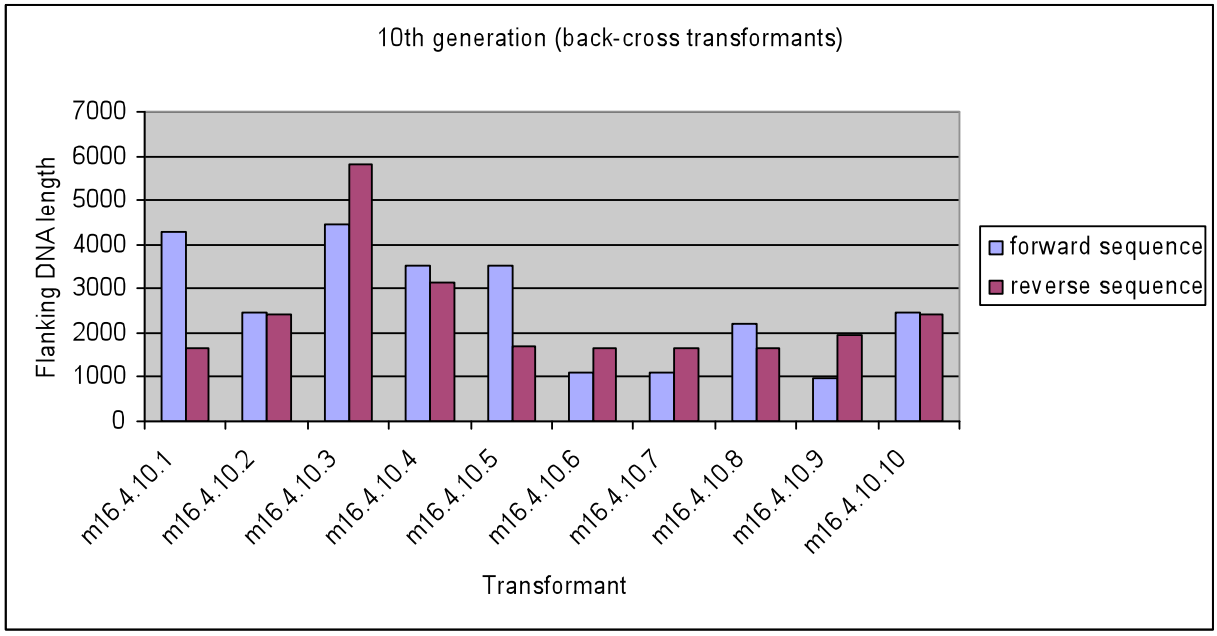


Figure 10. Transformants from *A. baylyi* 7021 were sequenced with different primers. The graph shows flanking DNA length of transformants among 10th generation on the left and right side of the *nptII*-gene. Maximum flanking DNA length in forward and reverse direction is equal to the recombination junction.

Sequencing with primer fw-8 (and fw-7) had resulted in several uncertain/doubtful fragments between 2701-3476 bases among all transformants.
Chromatograms from isolate M16.4.10.8 in fw-5, 7, 8 areas was difficult to study and the putative recombination junction was not clear, (Chromatograms are available in Sequencher-file in microbiology laboratory -UiT-IFA). Generally, during primer walking the sequences that were generated in some few points contained mismatches before they reach cross over junctions. Recombination junctions in all 10 back-cross transformants were identified (except back-cross transformant 8 in forward direction/right side of *nptII*-gene).

Average flanking DNA length

The sum of the length of flanking DNA in both forward and reverse direction of *nptII*- gene in the five initial and ten back-cross transformants respectively are: 62394 nucleotides and 50194 nucleotides. Average of flanking DNA length was assumed to compare 0^{th} generation with 10^{th} generation. The result indicates the difference in length with a high degree of confidence. Initial transformants has acquired a considerably longer DNA sequence from the donor (foreign DNA sequence) and only 3 crossover junctions were reached in only one direction/side of the *nptII*-gene. Figure 11 shows the average donor DNA length between the 0^{th} generation and 10^{th} generation compared with one another. Average DNA length submits in evidence the acquisition of a "longer" donor DNA in initial transormants, where recombination junctions were not identified in most of these. The comparison shows acquisition of shorter donor DNA sequence in backcross transformants where 19 recombination junctions were identified. Average flanking DNA length was about 12479 bases for initial isolates and about 5019 bases for back-transformants. Error bars indicate 95% confidence intervals ($\alpha = 0.05$).

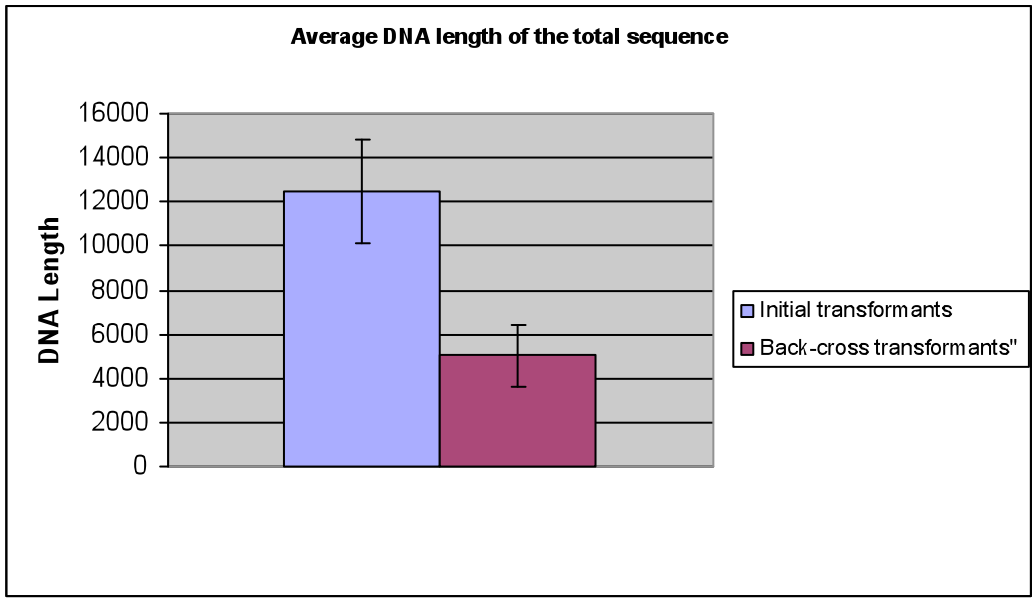


Figure 11. Average flanking DNA length for 5 initial *A. baylyi* 7021 transformants compared to 10th back-cross transformants**.** The sum of reverse and forward sequences (upstream and downstream sequences) of the *nptII*-gene for each tranformant was calculated and the average length of flanking DNA was obtained. The figure shows difference in the length of the flanking donor DNA (*A. baylyi* 62A1) between 0th generation where only 3 recombination junctions on one direction was identified in contrast with 10th generation in which 15 recombination junctions in both direction of *nptII*-gene was obtained. Error bars indicate 95% confidence intervals. In back cross transformants, average DNA length is equal to average length of the flanking donor sequence acquired by the recipient.

Total inserted size of donor DNA into recipient genome

Nucleotides/bases from the donor DNA flanked into genome of recipient DNA is between 12674 -14380 for four transformants within the five initial transformants. Only one transformant had a total insert donor sequence of 7864 bases. Insert size and length of the flanking sequences within back-cross transformants was between 2762 - 6643 in nine transformants. Only one transformant had a total insert length of 10291 bases. These results also indicates the total insert size of the donor and the sum of flanking sequences on the left and right side of the *nptII*-gene.

Discussion

Transformation as a major horizontal gene-transfer mechanism impacts genetic adaptation and evolution of prokaryotes. Chromosomal DNA transferred within species integrated by recombination depends on the systematic action of enzymes which joins and pair strands in addition to the necessity for a minimum degree of homology. The role of mismatch correction during recombination, by mismatch repair system as an important requirement for strand exchange and for the exogenous DNA to integrate is discussed below. In this study we investigate the fate of heterologous DNA (*A. baylyi* 62A1) acquired by the mutator strain *A. baylyi* ADP7021 as recipient and elimination of non-selected DNA sequence from the bacterial genome over ten generations during the process of introgression.

Suitability of methods

Isolation of genomic DNA by the Qiagen genomic DNA protocol, primer walking and DNA sequencing were the main methods in this project. Sequencing DNA from bacterial cell lysate was also examined:

DNA sequencing/ direct scanning PCR products

In this study, PCR-product from both genomic DNA and bacterial cell lysates was performed. The aim was sequencing DNA from hybrid transformants of *A. baylyi* to identify putative recombination regions. PCR products from bacterial genomic DNA and lysates were analyzed by gel electrophoresis. Flanking sequences in *A. baylyi* transformants on the left and right side of the inserted *nptII*- gene with primers downstream and upstream of the nptII-gene (forward- and reverse primes) were to be identified. This process was repeated several times with varieties of preparation methods and concentration of reagents. PCR-product analyzing of neither cell lysates nor genomic DNA by gel electrophoresis did not show migration of DNA on the gel except in one case, described in the section of results. This method was therefore not regarded suitable. The reason for lack of migration was unknown. One possible explanation is that fragments of the donor DNA has been transformed in different loci within the recipient, hence multi-recombination flanking regions on the genomic DNA of the transformants, might make the primers unsuitable to anneal with the template. The primers might anneal to nucleotide sequences other than the true target/template or dimeriz and anneal to themselves. The result of comparison by BLAST between the donor, 62A1 and the recipient ADP1, showed a very high homology between the donor and recipient DNA sequence in some regions (about 89-99%). In some region between 20-30 bases complete match (100 % similarity) and from this

to the next 30 bases, the similarity reduces suddenly to 29%. Another reason is possibly that primer pair LP1046, LPCRrv1- and rv2 which were used to sequencing PCR products, in the area 1000-3500 bp were not suitable for this region to anneal because of different annealing temperature (Tm ~ 78 ºC, 62 ºC respectively). This difference in temperature was not considered during the processes of performing.

Primer walk sequencing

Our goal in this work was to identify putative recombination junctions on the right and left ends of the *nptII* gene in 62A1/ADP7021 transformants per generation for 10 generation and five initial transformants. The results presented in this thesis are based on this method. The *nptII*-gene (Kanamycin resistance gene) consists of approximately 1220 bases. All transformants were sequenced with the first forward or reverse primer LP1046 and 2RP152 from a previous study (Ray, 2007) and sequences were compared against each other and had identical nucleotide sequences (matched) with a length between 800-850 bases.

The rest of the primers used in this study were designed using these sequences as basis for continuing the primer walking. Chromatograms were studied and sequences were compared and analyzed against ADP1- genomic sequence using the BLAST. It was not a simple matter to identify the differences in nucleotides sequences in all tranformants since we could only sequence a maximum length for each segment, about 600-850 bp in every step with a new primer. Some of the recombination junctions fell within regions where the nucleotide sequence (chromatogram) was easily readable using Sequencher without difficulties while some other regions needed to be more than one time sequenced and analyzed to be recognized. The reason why it was a problem to identify the donor sequence in some region of the nucleotide sequences is unknown. The quality of primers, the annealing temperature and the PCR program and preparing the samples, concentration of the template, Big Dye and primers were the same in each step. Modifying the concentration of DNA and primers did not affect the results, thereby new primers was designed for the specific transformants to continue primer walking the sequence. Contamination of the isolated genomic DNA was also excluded because using some other primers gave a clean (readable chromatogram) sequence as results for sequencing the same transformant. In doubtful cases of contamination, new DNA was isolated form the transformants. A positive control was used in most of the cases to excluding the possibility for personal or technical source of error. In some other cases, when a new achieved sequence was compared with BLAST and other sequences, it showed a different, single nucleotide, while the rest of the nucleotide sequence was not clear enough to study. After assembly of other sequences to compare the results in those “problem”-sequences the different nucleotide

(mismatch) that was shown in the specific donor DNA-sequence and using the BLAST to compare it against ADP1-genomic sequence, it was not a recombination junction. The reason why the base was different in some sequences compared to consensus sequence is uncertain. Viewing appendix 1- contig for individual transformants, there are 1-2 gaps between some sequences which indicates the difficulty of primer walking in those regions using different primers. Nevertheless this method gave the opportunity to achieve important results in this study.

Analysis of the flanking DNA length integrated in the transformants

Results from this study submit the hypothesis that non-selected DNA is eliminated from the recipient genome through introgression. Repeated back crossing of the *A. baylyi* hybrid transformants with ADP 7021 compared to initial transformants confirm that in this experiment with 15 transformants, the length of donor DNA in the 5 initial transformants is significantly larger than 10^{th} generation back-cross transformants. There was significant difference in the average flanking DNA length of fragments integrated by 0^{th} generation and 10^{th} back-cross generation. Some of the flanking regions were identical and had the same nucleotide position. We found that the donor fragments were smaller in 10^{th} generation back-cross isolates and most of them did not extend past base 1671 in reverse – and 2500 in forward direction of the *nptII*-gene compared to more that 12000 bases among 0^{th} generation. There are a number of studies indicating the importance of the length of flanking homology in relation to transformation frequency (De Vries *et al*., 2002, Harries-Warrick *et al.,* 1978, Majeswski *et al*., 1999, Watt *et al*., 1985). The size of donor DNA fragment in a recipient has been discussed in studies of Bacillus and it was suggested that a longer donor fragment may potentially result in greater stability to the donor-recipient heteroduplex under condition of transformation (Zawadzki *et al*., 1995). They concluded that longer donor fragment may also saturate the action of mismatch repair enzymes since their action is believed to be directed on the donor-recipient heteroduplex under the process of recombination. Transformation with genomic DNA in their study resulted in integration of longer donor fragment comparing to PCR-amplified DNA. Hence assuming that longer donor fragments from a genomic donor DNA, has effect on transformation in which it yields longer integrated fragments into a recipient. In this study a mutator type of *A. baylyi* was used, the long flanking donor sequence into the recipient seems to be a consequence of deletion of *mutS*-gene in *A. baylyi* ADP7021, because the mismatch correction is deficient. Between the back cross transformants, the results indicate that among 10 transformants in this study, non selected heterologous sequence during introgression is eliminated from the bacterial

genome. Longer flanking DNA sequence in the 5 initial- compared to back cross transformants confirm the effect of introgression.

Divergence between recipient and donor

Sequence divergence and mismatch repair system is known as important barriers to suppress recombination processes between bacteria. MutS protein can recognize seven of eight possible base pair mismatches and inactivation of *mutS* or *mutL* genes increases recombination between non identical sequences (Denamur and Matic, 2006). Analyses from earlier studies have shown that bacterial heterologus recombination can introduce a donor DNA in the recipient from extremely divergent taxa (Majewski and Cohan, 1998; Matic *et al.,* 1995). Bacteria can undergo homologous recombination with organisms that differ by at least 25% in DNA sequence (Duncan *et al*., 1989, Zhou and Spratt, 1992, Cohan, 2002). Increased sequence divergence reduces recombination among species. Earlier results from one study by Zawadzki *et al* (1995) investigated the relationship between sexual isolation and sequence divergence and that it may be affected by the length of donor DNA fragments presented to the recipient cell (in Bacillus transformations). It is mentioned that one condition that might affect the sensitivity to transformation to sequence divergence is the length of DNA presented in recipient cell because longer fragment from a mismatch DNA may reduce the effectiveness of mismatch repair enzymes. Transformation of flanking donor DNA might results in integration of short DNA segments as a single fragments or longer segments may integrate as multiple discontinuously fragments. The potential for integration is greater when the DNA from a divergent donor is presented in smaller fragments according to earlier studies. Factors reducing recombination between divergent bacteria described by Cohan (2002), in summary are as follows: ecological differences (physical distance between species), behavioural changes in micro organisms (bacteria may become competence to gene exchange under specific circumstances, as starving), DNA entry prevention (when transformation system of a recipient and a donor sequence do not give admission to foreign DNA entry), restriction system's impact, as endonuclease activity and effectiveness (cleavage of DNA at specific target sequences), the degree of integration of donor DNA (the importance of the stability of recombined DNA into the recipient genome by an identical segment), activity of mismatch repair enzymes (in which are responsible to resistance to recombination between species and the incompatibility and deficiency of foreign alleles transferred by recombination into a recipient). Interspecific recombination is also related to intrinsic competence (IC) for transformation within strains of a species. In other words, depended on the level of IC within a natural population, transformation rate might vary and will

be affected (lower or higher) for the strains compared to intraspecific transformation. Modifying this level will change the capacity of interspecific recombination (Cohan 2002; Cohan *et al*., 1991). Thus the activity of mismatch repair enzymes, sequences divergence, and species' sensitivity to sequences divergence has been considered as important factors affecting the rate of transformation and diversity hence sexual isolation among species. Transformants in this study performed by the previous study of Ray (2007) shows that in the process of transformation large DNA sequences from the 25% divergent donor have been recombined into the recipient. In studies of *Acinetobacter* sp. strain ADP1 (Young and Ornston, 2001), inactivation of *mutS* resulted in 3- to 17-fold increase in transformation efficiencies with a donor sequence that were 8-20 % divergent to the strain ADP1. The heterologous donor strain *A.* sp. 62A1 is approximated 25% divergent from *A. baylyi* at the DNA level. The result may be due to the inactivation of *mutS* gene in addition of the intrinsic competence of *Acinetobacter* for transformation.

Impact of horizontal gene transfer (HGT) on bacterial diversity

HGT has shown to be one of the main processes and forces for movement of genes from one bacterium to another. Gene transfer and exchange may allow bacteria to adapt new properties. It is believed that some of the divergent sequences in species are brought about by gene acquisition and adaptation unrelated with cell division (Lawrence, 2005). Based on studies of strains such as *Streptococcus suis* and *Helicobacter pylori* it is considered that variation in restriction-modification system may be a consequence of horizontal gene transfer. These strains have shown different, unique heterolougous restriction modification systems compared to other strains within the species and are believed to be acquired through horizontal gene transfer. Mutations are, in addition found to have impact on heteroduplex formation between divergent DNA strands in which increases the sensitivity of bacteria to sequence divergence as in the case of mutation in certain genes in *B. subtilis*. Mismatch repair genes variation appears to be generated through horizontal gene transfer, as example in wild isolates of *E. coli* which contains widely divergent sequences. Acquisition of specific genes from other species by horizontal gene transfer will lead to modification and variation in the genome of at recipient among a population of micro organisms. There is a number of evidence to confirm that HGT is a way of exchange of genetic material which plays a major role in the diversity between or within species. Increasing antibiotic resistance in human pathogens is a phenomena derived from this important type of gene movement in environment.

Benefits of the DNA uptake in bacteria

Uptake of extracellular DNA has been also proposed to be used for nutritional purpose by the recipient where it is depended on homology between bases (base pairing), to increase the ability to survive and reproduce. It has also been discussed that DNA uptake by a recipient cell occurs to be used in repair of DNA lesions. Benefits of flanking donor segments has been considered against its general deleterious effects in the genome of recipient bacteria by horizontal gene transfer (Thomas *et al*., 2005). These aspects are challenging in studies of bacterial horizontal gene transfer. Recombination between DNA sequences and genome rearrangements may by chance be beneficial to bacteria. It is believed that, generally, most of the acquired genes have deleterious effects in the chromosome of a recipient bacterium but it is speculated that beneficial effects are favoured by bacteria. Adaptations may allow transfer of a segment of value for the recipient. Despite the benefits in some cases, there is a fitness cost to uptake of foreign DNA and recombination might reduce the fitness in bacteria as in studies of *E. coli* and *Salmonella* discussed by Cohan (2002). Recombination by transformation is able to transfer long segments up to several thousand base pairs. The balance between benefit of excessive acquisition and the cost may be reached by high variability in the size of the fragments (Cohan, 2002). Transforming fragments brings a series of maladaptive genes which is necessary to flank the useful adaptive genes. To my knowledge there is no study investigating the consequence of elimination of foreign DNA incorporated into a recipient DNA in relation to horizontal gene transformation. It would be interesting to study the effect and consequence of these long flanking DNA fragments in *A. baylyi* transformants genome related to its functional changes and pathogenesis. In the case of Bacillus, it has been shown that the size of transferred fragment by recombination was up to around 10 kb (Zawadzki *et al,* 1995). In this study we found that foreign DNA from a 25 % divergent strain (donor DNA containing kanamycin resistance gene) is transferred from a range of several hundred base pairs to fragments of over 13 kb and has been recombined in *A. baylyi* 7021 hybrid transformants. Recombination between divergent species involves incorporation of genes of a whole operon segment and adaptation by a recipient (Cohan, 2002). The reason why recombination in this case involves integration of longer flanking fragments might be related to the defect in the mismatch repair system of the recipient *A. baylyi* and or the sequences acquired may be beneficial to the recipient. In addition, *A. baylyi* known as a high competent bacterium for transformation seems to have the ability to recombine with longer fragments of the foreign DNA. The defect in mismatch repair system, (*mut*S deletion) is suggested as a reason to acquisition of the large amount DNA, since

transformation rate increases when the mismatch rapier system is defect and because of the lack of enzymatic activity under the formation of heteroduplex.

Conclusion

In this study identifying the recombination junctions and analyzing the flanking DNA size in the *A. baylyi* 7021 transformants results in achieving important aspects to further works. The difference between the size of flanking sequence in 0^{th} transformants and 10^{th} back transformants was shown under the process of introgression. The size of flanking DNA was smaller in back-cross transformants comparing with the initial transformants. Foreign DNA is eliminated by back crossing of the transformants but because of the limitation (only ten back-cross isolates) of this study it is not possible to make a direct conclusion about the size of the foreign DNA within the 10^{th} generation, in other words, among the first backcross (isolate 1) and tenth backcross transformants (isolate 10). Based on these results, the integration of shorter donor DNA into the recipient might have been balanced with loss of "non-selected" foreign sequence among 10^{th} back transformants under the process of introgression. In transformation with genomic DNA, large donor molecules may lead to integration of long donor segments while smaller segment from most divergent donor has higher potential to integrate (Zawadzki *et al*., 1995). Heterologus donor DNA can be recombined by "anchors" (small homology sequences) between an introducing exogenous DNA and a competent recipient. Anchor region facilitate recombination and insertion of foreign, divergent DNA sequences. It is believed that introgression of genes can occur by the help of anchors under condition of horizontal gene transfer where it involves gene acquisition and gene loss. Initial transformants had acquired much longer donor sequence this is a new aspect comparing to studies of *Bacillus* and *E. coli*. The fact that defect in mismatch repair enzymes increases the rate of transformation of exogenous DNA between species and insert size of the donor sequence is discussed in several studies as mentioned before. The effect of mismatch repair system and *mutS*-gene deletion is questionable in backcross generation where shorter DNA sequence was flanked into the recipient ADP7021 comparing to initial transformants. During introgression, *mutS* deficient might have weak effect and or have no influence in the flanking size of donor DNA. If this is the case, thus it will impact genome evolution in bacteria and would contribute to exchange to genetic information between non related species.

References

Andersson, S. G. E. and Kurland C. G. (1998). Reductive evolution of resident genomes, Trends in Microbial. 6, 263-268.

Arber, W. (1995). The generation of variation in bacteria genomes. J. Mol Evol. 40, 7-12.

Barbe, V., Vallenet D., Fonknechten, N., Kreimeyer, A., Oztas, S., Labarre L., Cruveiller, S., Robert, C., Duprat, S., Wincker, P., Ornston, L.N., Weissenbach, J., Marlière, P., Cohen, G.N., Médigue, C., (2004). Unique features revealed by the genome sequence of *Acinetobacter* sp. ADP1, a versatile naturally transformation competent bacterium. Nucleic Acid Res. 32, 5766-5779.

Brown, E.V., Mammel, M.K.J., LeClerc, E. and Cebula T. (2003). Limited boundaries for extensive horizontal gene transfer among *Salmonella* pathogens**.** Proc. Natl. Acad. Sci USA 100, 15676–15681.

Brown, E.W., LeClerc, E. J., Li, B., Payne, W. L., and Cebula, T. A. (2001). Evidence for horizontal transfer of *mutS* alleles among naturally occurring *Escherichia coli* strains. J. Bacteriol., 183, 1631-1644.

Buchan, A. and Ornston, N.L. (2005). When coupled to natural transformation in *Acinetobacter* sp. strain ADP1, PCR mutagenesis is made less random by mismatch repair. Appl. Environ. Microbiol., 71, 7610–7612.

Cohan, F. M. (2002). Sexual isolation and speciation in bacteria. Genetics 116, 359-370.

Cohan, F. M., Roberts, M. S., King E. C. (1991). The potential for genetic exchange by transformation within a natural population of Bacillus subtilis. Evolution, 45, 1393-1421.

De Vries, J. and Wackernagel, W. (2002). Integration of foreign DNA during natural transformation of *Acinetobacter* sp. by homology-facilitated illegitimate recombination. Proc. Natl. Acad. Sci. USA., 99, 2094-2099.

Denamur, E. and Matic I. (2006). Evolution of mutation rates in bacteria. Molecular biol. 60, 820-827

Denamur, E., Lecointre, G., Darlu, P., Tenaillon, O., Acquaviva, C., Sayada C., Sunjevaric, I. Rothstein, R., Elion, J., Taddei, F. (2000). Evolutionary implications of the frequent horizontal transfer of mismatch repair genes. Cell 103, 711-721.

Duncan K. E., Istock C. A., Graham J. B., Ferguson N. (1989). Genetic exchange between Bacillus subtilis and Bacillus licheniformis: Variable hybrid stability and the nature of bacterial species. Evol., 43, 1585-1609.

Dutta, C. and Pan A. (2002). Horizontal gene transfer and bacterial diversity. J. Biosci. 27, 27-33.

Finkel, S. E. and Kolter, R. (2001). DNA as a nutrient: novel role for bacterial competence gene homologes. J. Bacteriol. 183, 6288-6293.

Garcia-Vallve, S., Janssen, P. J. and Ouzounis,C. A. (2003). Genetic variation
HGT-DB: a database of putative horizontally transferred genes in prokaryotic complete genomes. Nucleic Acids Research, 31, 187-189.

Gomez, M. J. and Neyfakh, A. (2006). Genes involved in intrinsic antibiotic resistance of *Acinetobacter baylyi.* Antimic. Agents and Chemother. 50, 3562-3567.

Gore, J. M. F., Ornston, A. R. and Nicholas, L., (2006). Deletion mutations caused by DNA strand slippage in *Acinetobacter baylyi*. Environ. Microbiol. 5239–5245.

Harris-Warrick, R. M. and Lederberg, J. (1978). Interspecies transformation in bacillus: sequence heterology as the major barrier. J. Bacteriol. 133, 1237-1245.

Lawrence, J. G. (2005). Horizontal and vertical gene transfer: the life history of pathogens. In Russel W, Herwals H. eds, Concepts in bacterial virulence, Basel, Karger, pp 255-271.

Lorenz, M. G. and Wackernagel, W. (1994). Bacterial gene transfer by natural genetic transformation in the environment. Microbiol. Mol. Biol. Rev. 58, 563-602.

Majewski, J. and Cohan F. M. (1999). DNA sequence similarity requirements for interspesific recombination in Bacillus. Genetics, 153, 1525-1533.

Majewski, J. and Cohan F.M., (1998). The effect of mismatch repair and heteroduplex formation on sexual isolation in Bacillus. Genetics 148, 13-18.

Matic, I., Rayssiguier, C., and Radman M. (1995). Interspecies gene exchange in bacteria: The role of SOS and mismatch repair system in evolution of species. Cell 80, 507-515.

Matic, I., Taddei, F., and Radman , M. (1996). Gentic barriers among bacteria. Trends Microbiol. 4, 69-73.

Nielsen, K. M., Van Elsas, J. D. and Smalla, K. (2000). Transformation of *Acinetobacter* sp. Strain BD413 (pFG4ΔnptII) with transgenic plant DNA in soil microsoms and effects of Kanamycin on selection of tranformants. Appl. and Environ. Microbiol., 66, 1237-1242.

Ochman, H., Lawrence, J. G. and Groisman, E. (2000). Lateral and oblique gene transfer and the nature of bacterial innovation. Nature 405, 299-304.

Ray, J. L. (2007). Natural transformation with chromosomal DNA in bacteria: occurrence and determinative factors. Ph.D. thesis, Department of Pharmacy, University of Tromsø, Norway.

Salyers, A. A. (1995). Antibiotics resistance transfer in the mammalian intestinal tract: implications for human health, food safety and biotechnology. New York, Springer.

Simpson, D. J. and Dawson, L.F., Fry, J. C., Rogers, H. J., Day, M. (2007). Thematic issue on horizontal gene transfer; Influence of flanking homology and insert size on the transformation frequency of *Acinetobacter* baylyi BD413. Environ. Biosafety Res. 6, 1-15.

Singer, J. and Finnerty, W. R. (1984). Insertional specificity of transposon Tn5 in *Acinetobacter* sp. J. Bacteriol. 157, 607–611.

Taylor, W. H. and Juni, E. (1961). Pathways for biosynthesis of a bacterial capsular polysaccharide I. Characterization of the organism and polysaccharide. J. Bacteriol. 81, 688-693.

Thomas, C. M. Nielsen, K. M. (2005). Mechanisms of and barriers to horizontal gene transfer between bacteria. Nature Microbiol. Rev. 3, 711-921.

Tønjum, T., Håvarstein, L. S., Koomey, M. and Seeberg, E. (2004). Transformation and DNA repair: linkage by DNA recombination. Trends in Microb. 12, 1-4.

Townsend, J. P., Nielsen K. M., Fisher D.S., Hartl D.L. (2003). Horizontal acquisition of divergent chromosomal DNA in bacteria: effects of mutator phenotypes. Genetics 164**,** 13-21.

Trun, N. and Trempy, J. (2004). Fundamental Bacterial Genetics. Malden, Mass. Blackwell Sience.

Watt, V. M., Ingles, C. J., Urdea, M. S., Rutter, W. J. (1985). Homology requirements for recombination in *Eschericha coli*. Proc. Natl. Acad. Sci. USA. 82, 4768-4772.

Young, D. M. and Ornston, L. N. (2001). Functions of the mismatch repair gene *mutS* from *Acinetobacter* sp. Strain ADP1. J. Bacteriol. 183, 6822-6831.

Zawadzki, P. and Cohan F. M. (1995). The size and continuity of DNA segments integrated in Bacillus transformation. Genetics 141, 1231-1243.

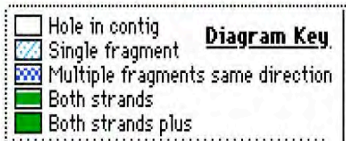
Zawadzki, P., Roberts, M. S. and Cohan F. M. (1995). The log-linear relationship between sexual isolation and sequence divergence in Bacillus transformation is robust. Genetics 140, 917-932.

Zhou, J. and Spratt B. G. (1992). Sequence diversity within the *argF, fbp* and *recA* genes of natural isolates of *Neisseria meningitidis*: interspecies recombination within the *argF* gene. Mol. Microbiol.6, 2135-2146.

Appendix 1- contig overview - Assembled contigs using Sequencher™ version 4.1.4

Assembly of DNA sequences on left and right side of the *nptII*-gene for all transformants resulting from primerwalk sequencing.

Key diagram:



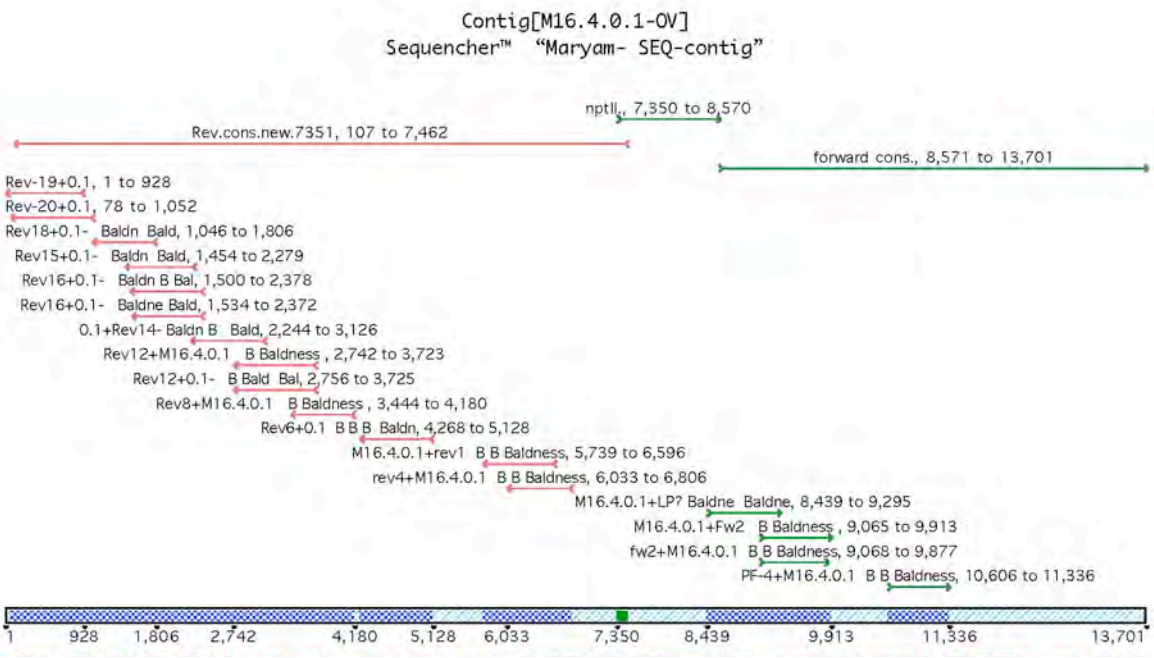


Figure A1. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.0.1
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping

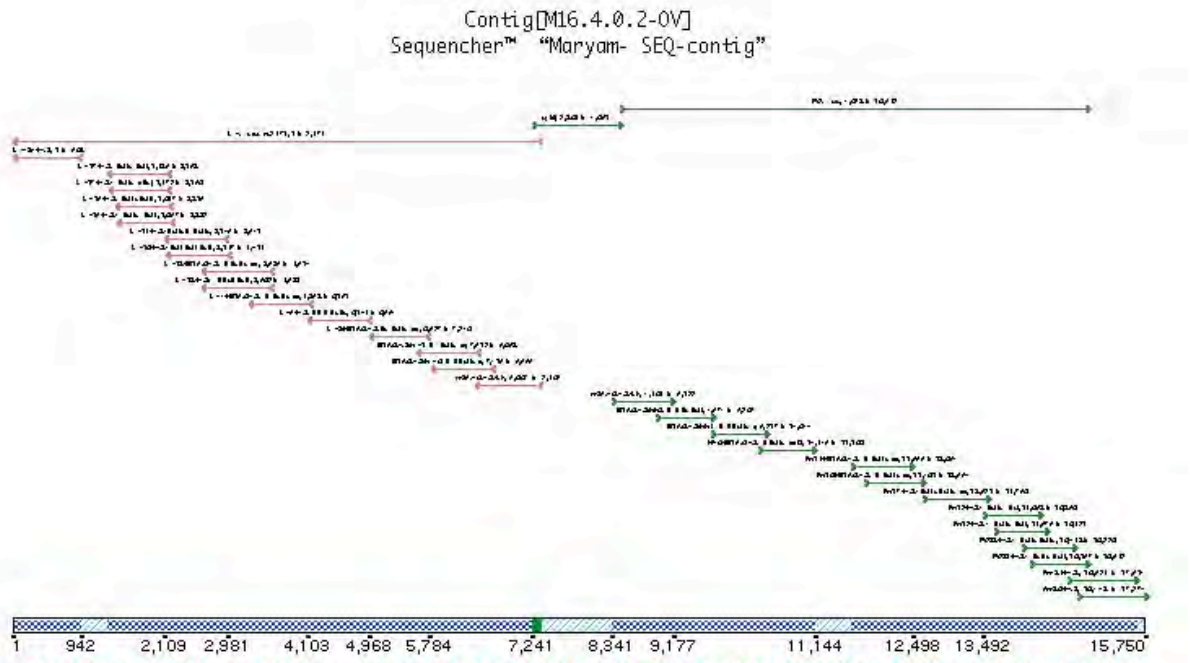


Figure A2. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.0.2
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping

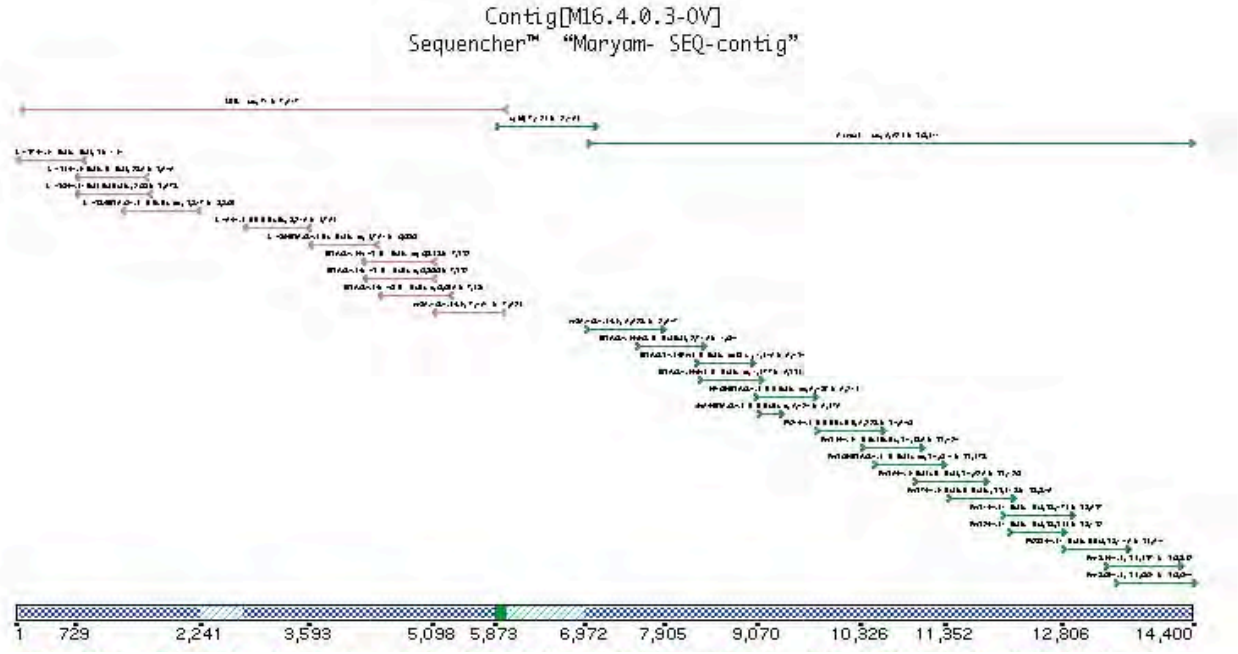


Figure A3. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.0.3
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary

line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping

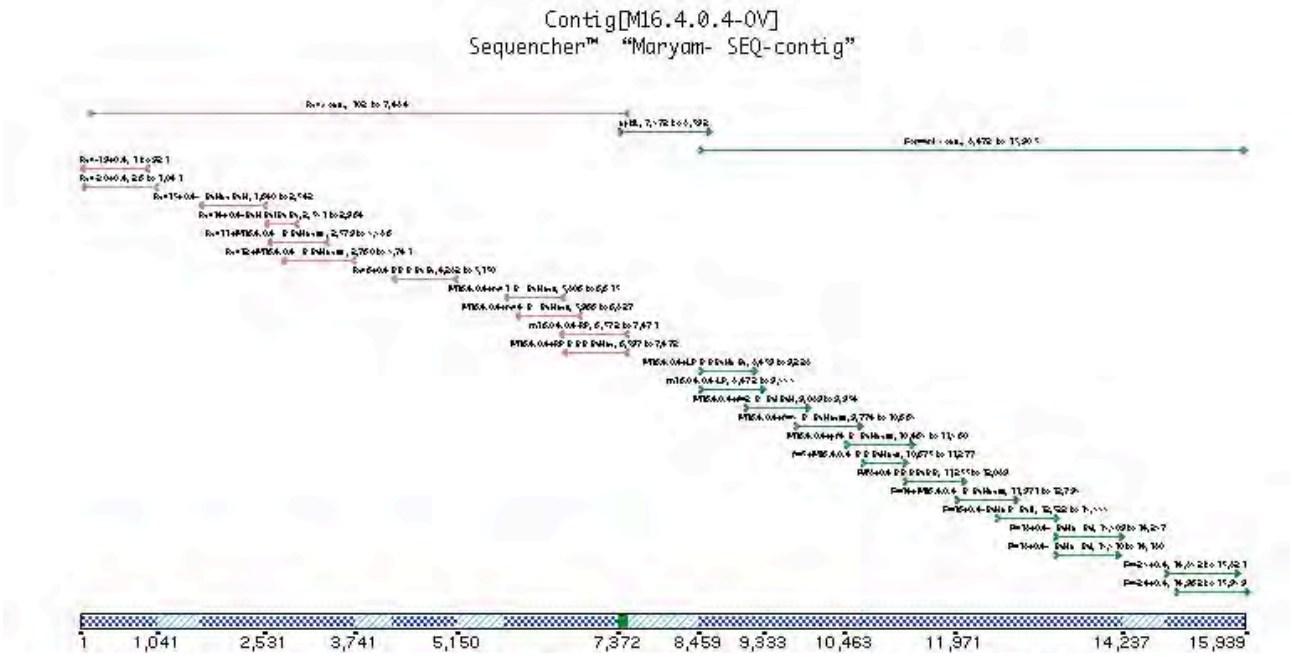


Figure A4. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.0.4
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping

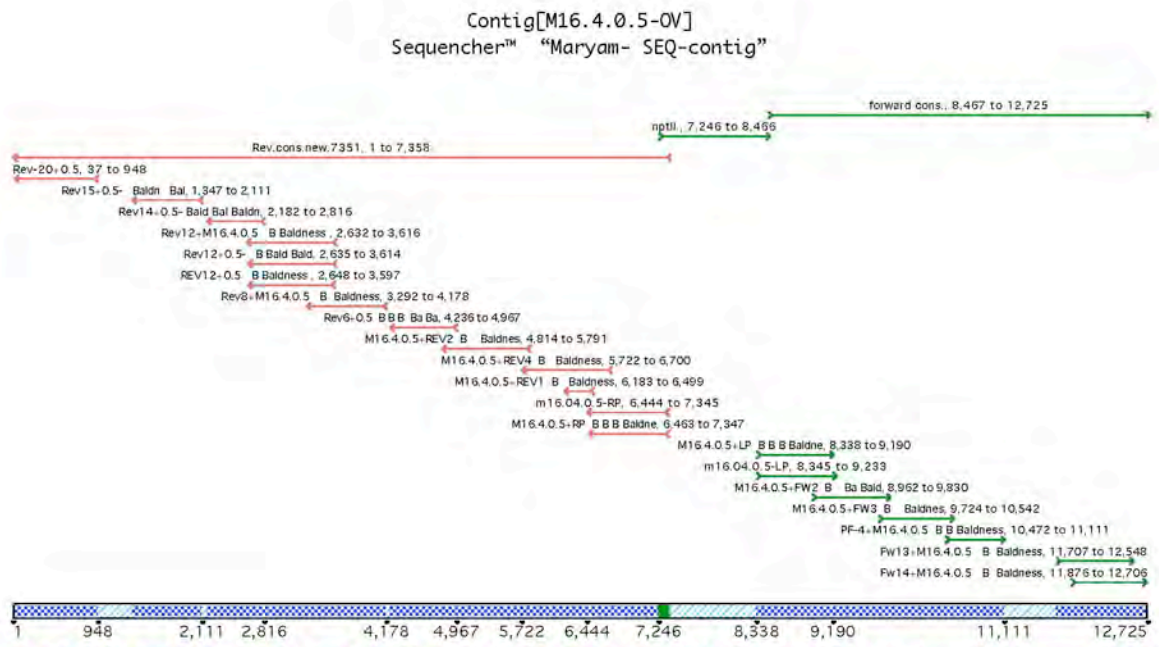


Figure A5. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.0.5
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping

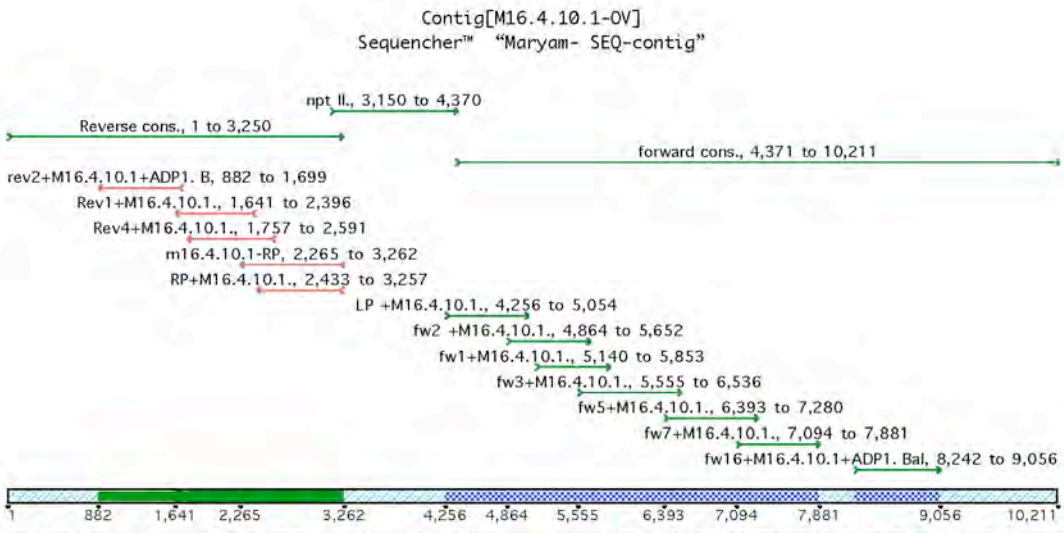


Figure A6. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.10.1
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping.

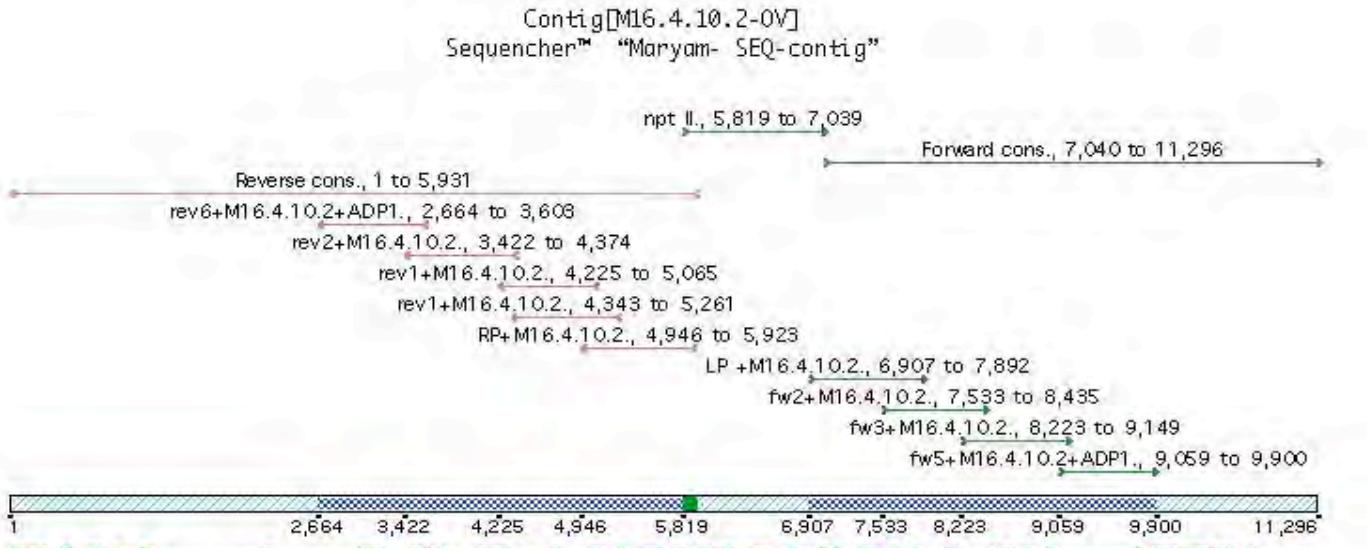


Figure A7. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.10.2
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping.

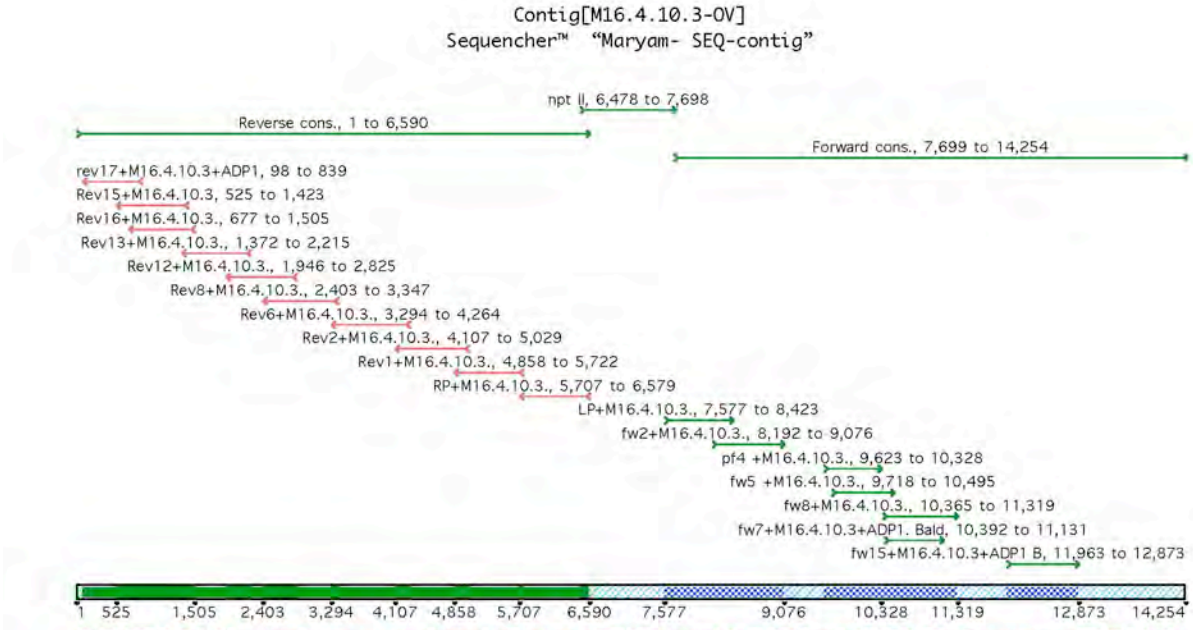


Figure A8. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.10.3
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary

line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping.

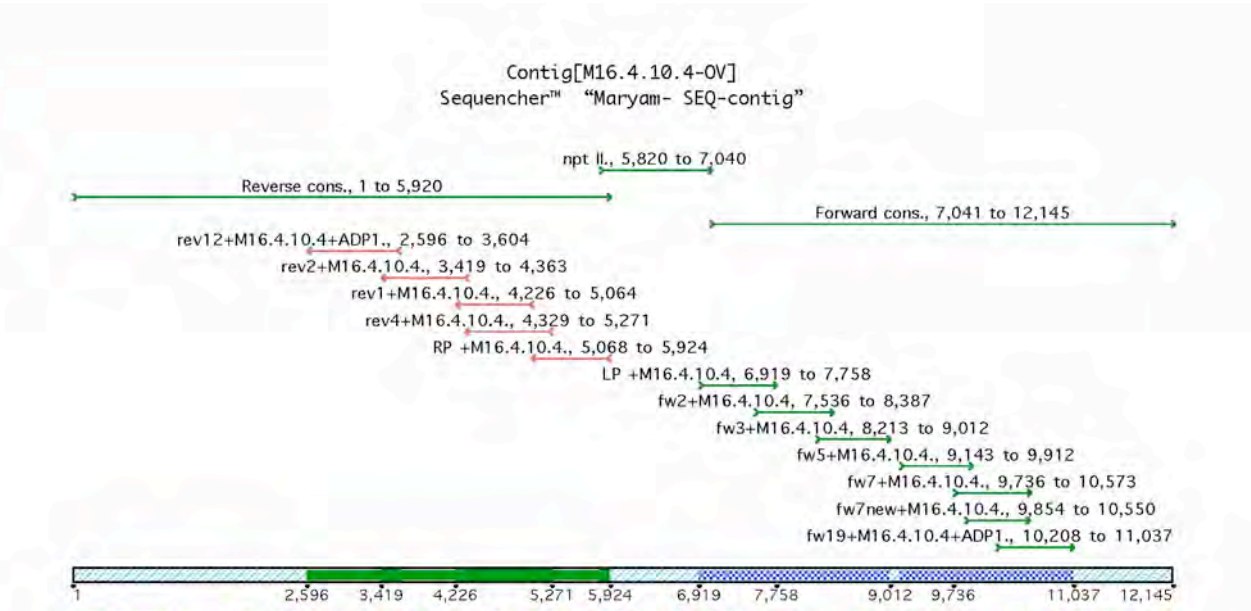


Figure A9. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.10.4
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping.

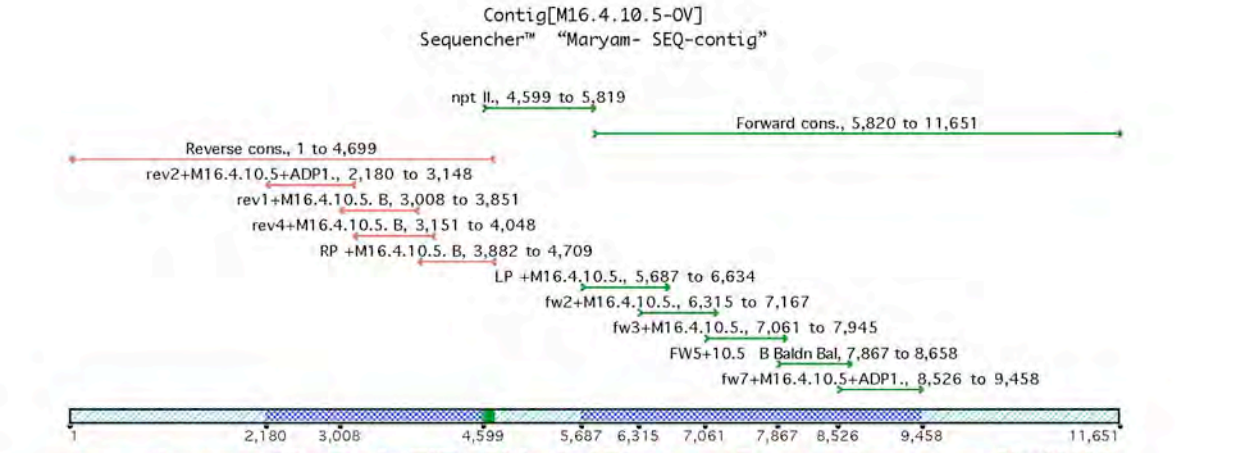


Figure A10. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.10.5
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary

line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping.

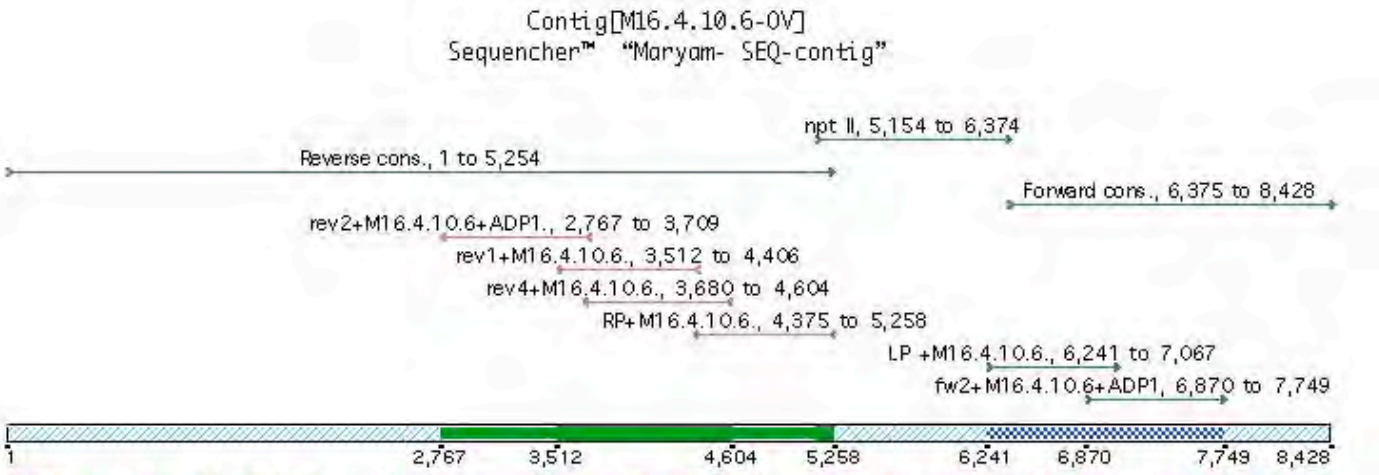


Figure A11. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.10.6
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping.

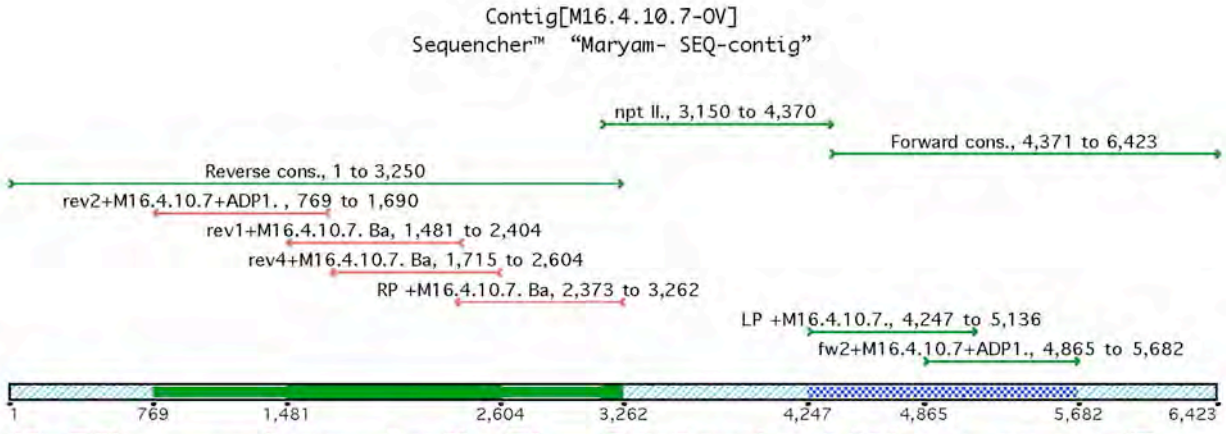


Figure A12. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.10.7
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping.

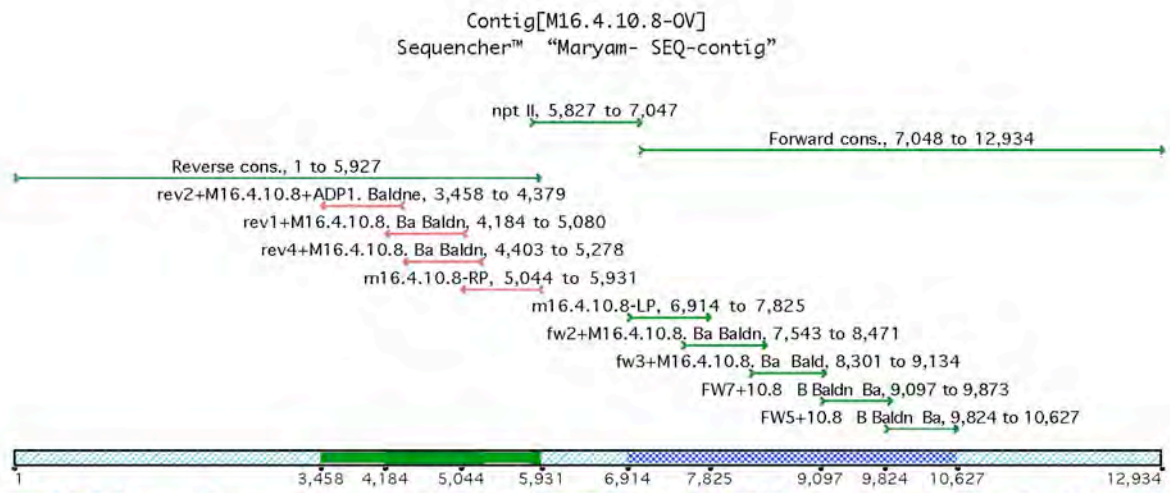


Figure A13. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.10.8
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping.

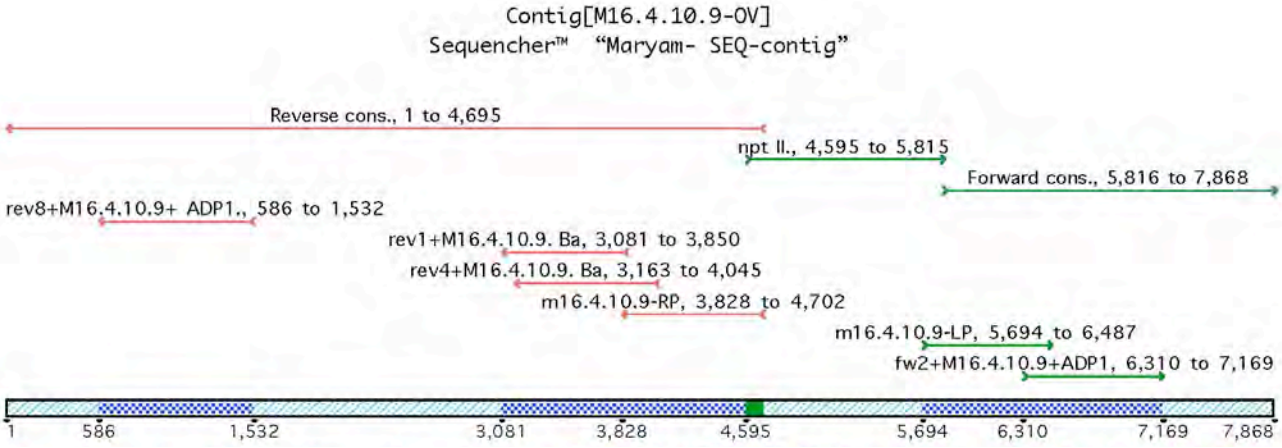


Figure A14. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.10.9
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping.

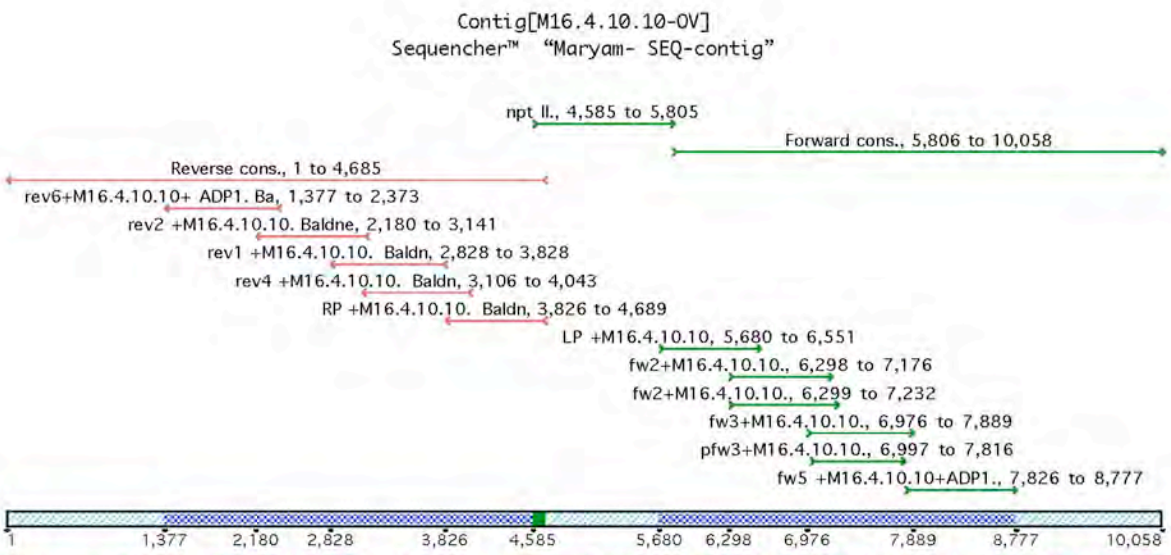


Figure A15. Contig of the primer walk sequence of *A. baylyi* transformant M16.4.10.10
The forward and reverse consensuses flanking *nptII*-gene are shown in the figure. The assembly of sequences is demonstrated with primers toward the sequence generated. The contig summary line indicates: lightly strand, single fragment, shaded-multiple fragments have the same direction, solid shading strands are overlapping.

Appendix 2. DNA quantification

DNA concentration measured by The NanoDrop® ND-1000

ng/ul: sample concentration in ng/ul based on absorbance at 260 nm and the selected analysis constant. The spectrophotometer can be used to quantify purified DNA at OD_{260}. DNA at 50μg/ml has an OD_{260} of 1.0.

Data in table 18 is from the first batch of DNA isolated by Qiagen at the start of this study. Samples with a concentration greater than 600 ng/ul were adjusted to approximately 300 - 500 ng/ul.

Table 18. Nanodrop Data 14.03.2007 & 26.03.2007 & 02.04.07

Sample ID	ng/uL	A 260	260/230	260/280	Const.
M16.4.0.1	**357.58**	**7.152**	**2.33**	**1.89**	**50**
M16.4.0.2	**373.11**	**7.462**	**2.23**	**1.79**	**50**
M16.4.0.3	**460.04**	**9.201**	**2.32**	**1.90**	**50**
M16.4.0.4	**470.10**	**9.402**	**2.30**	**1.84**	**50**
M16.4.0.5	**420.97**	**8.419**	**2.32**	**1.79**	**50**
M16.4.10.1	**217.49**	**4.350**	**2.38**	**1.89**	**50**
M16.4.10.2	**859.43**	**17.189**	**2.35**	**1.91**	**50**
M16.4.10.3	**500.49**	**10.010**	**2.31**	**1.86**	**50**
M16.4.10.4	**563.23**	**11.265**	**2.31**	**1.85**	**50**
M16.4.10.5	**510.96**	**10.219**	**2.33**	**1.86**	**50**
M16.4.10.6	**1006.**	**20.128**	**2.32**	**1.91**	**50**
M16.4.10.7	**593.53**	**11.871**	**2.34**	**1.86**	**50**
M16.4.10.8	**650.48**	**13.010**	**2.21**	**1.85**	**50**
M16.4.10.9	**481.80**	**9.636**	**2.16**	**1.87**	**50**
M16.4.10.10	**847.82**	**16.956**	**2.15**	**1.8**	**50**

"The reading at 280 nm determines the amount of protein in a sample. The ratio between the readings at 260 nm and 280 nm (OD260/OD280) provides an estimate of the purity of the nucleic acid. Pure preparations of DNA and RNA have OD260/OD280 of 1.8 and 2.0 respectively. If there is contamination with proteins or phenol, the OD260/OD280 will be significantly less, and accurate quantitation of the amount of nucleic acid will not be possible. The spectrophotometer can be used to quantify purified DNA at OD_{260}. Strong absorbance around 230 nm can indicate that organic compounds or salts are present in the purified nucleic acid. A ratio between the readings at 260 nm and 230 nm (OD260/OD230) can help to evaluate the level of salt carryover in the purified nucleic acid. As a guideline, the OD260/OD280 is best if greater than 1.5."