UPTEC X 05 034
JUL 2005

ISSN 1401-2138

PETER LIND

# Genome reduction in *Salmonella enterica* by selection of loss of mini-transposon Tn10dTetKan

Master's degree project

UPPSALA UNIVERSITET

# Molecular Biotechnology Programme

Uppsala University School of Engineering

| **UPTEC X 05 034** | **Date of issue 2005-07** |
|---|---|


Author

**Peter Lind**


Title (English)

**Genome reduction in *Salmonella enterica* by selection of loss of mini-transposon Tn10dTetKan**

Title (Swedish)


Abstract

By experimentally evolving a reduced genome, knowledge about which genes are needed in a minimal genome and which mechanisms are involved in gene loss can be gained. A bacterium with a considerably reduced genome can be used for gene gain studies, as a less complex model organism and as a starting point in biotechnological applications. In this project a mini-Tn10 with two resistance markers is used to introduce large deletions in *Salmonella enterica*. Seventeen candidate deletions were found in two different chromosomal locations.



Keywords

*Salmonella enterica*, minimal gene set, genome reduction, deletion mutant



Supervisors

**Dan Andersson & Sanna Koskiniemi**
**Department of Medical Biochemistry and Microbiology, Uppsala University**

Scientific reviewer

**Linus Sandegren**
**Department of Medical Biochemistry and Microbiology, Uppsala University**



| Project name | Sponsors |
|---|---|
| Language **English** | Security |
| **ISSN 1401-2138** | Classification |
| Supplementary bibliographical information | Pages **26** |



**Biology Education Centre** Biomedical Center Husargatan 3 Uppsala
Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 555217

# Genome reduction in *Salmonella enterica* by selection of loss of mini-transposon Tn10dTetKan

Peter Lind

**Populärvetenskaplig sammanfattning**

Den genetiska informationen som överförs från en generation till nästa är kodad i sekvensen hos DNA, vare sig det gäller bakterier eller människor. Evolutionen innebär att då livsmiljön förändras kommer längden och sekvensen hos DNA efter hand att förändras då de individer som är bäst anpassade till omgivningen ökar i antal. Bakterier fortplantar sig så snabbt att man kan studera evolution i realtid. Trots att bakterier har mycket färre gener än till exempel människor så behövs de flesta inte för att de ska överleva i en stabil och näringsrik miljö.

Här används bakterien *Salmonella enterica* för att undersöka hur det går till när bakterier förlorar gener. Målet är att ta bort så många gener som möjligt för att se vilka och hur många gener som behövs för att bygga upp en fungerande cell. I 1 milliliter kan det finnas en miljard bakterier, så för att hitta de som förlorat gener selekteras för dem som förlorat resistens mot två olika antibiotika. Sedan undersöks vilka gener som förlorats och exakt vilken del av DNA molekylen som försvunnit. Bakterier med få gener skulle kunna användas för att se hur det går till när nya gener kommer till, som en enklare modellorganism samt som en utgångspunkt för applikationer inom biotekniken.

**Examensarbete 20 p i Molekylär bioteknik**

**Uppsala universitet juli 2005**

## Contents Page

# 1 Introduction and background

## *1.1 Salmonella enterica*

*Salmonella enterica* are gram-negative bacteria associated with gastro-intestinal disease. There are three main serovars with 95-99% DNA identity, Typhi, Enteritidis and Typhimurium.[1]

*Salmonella enterica* serovar Typhi cause typhoid fever, a serious disease with 17 million cases and 600,000 deaths annually almost exclusively in less developed countries in Asia and Africa. The pathogen is usually spread through contaminated water or food and it is only known to infect humans, so there are no known animal reservoirs.[2]

Serovar Enteritidis is a common cause of food-poisoning, often transmitted through chicken, and the symptoms include week-long diarrhea, abdominal cramps, vomiting and nausea and can be fatal if not treated with antibiotics.[1]

*Salmonella enterica* serovar Typhimurium can also cause food-poisoning, with symptoms similar to those of serovar Enteritidis. As the name implies it also cause typhi-like disease in mice. It has commonly been used as a mouse model for typhoid fever as well as a general bacterial model for a wide variety of biochemical and genetic experiments. The whole genome of the principal laboratory strain *Salmonella enterica* serovar Typhimurium LT2 (in this work referred to as *S. typhimurium*) has been sequenced. The chromosome of *S. typhimurium* is 4,857,432 bp long and the strain harbors a plasmid, pSLT (93,939 bp).[3]

## 1.2 The bacterial genome

The genetic information that is transferred from one generation to the next is encoded in DNA (deoxyribonucleic acid). In most bacteria the majority of the genes are encoded on a single circular DNA molecule, the bacterial chromosome. In general, it contains house-keeping genes, as well as most of the species-specific genes. The genome is the sum of all genetic information in the cell, which includes not only the chromosome but also plasmids and phages, which are independent replicons. Plasmids often contain genes allowing conjugation and genes conferring resistance to antibiotics.

It is important to realize that the bacterial genome is not a static entity. The size of the genome can increase by horizontal gene transfer and duplications and genes are lost due to deletions (Figure 1). This forms the basis of the generation of genetic variability crucial in adaptation to novel environments. There is a deletional bias, meaning that deletions are more common than insertions, so that genes that have been inactivated by mutations or genes that do not confer a selective advantage are often lost in the course of evolution.[7]

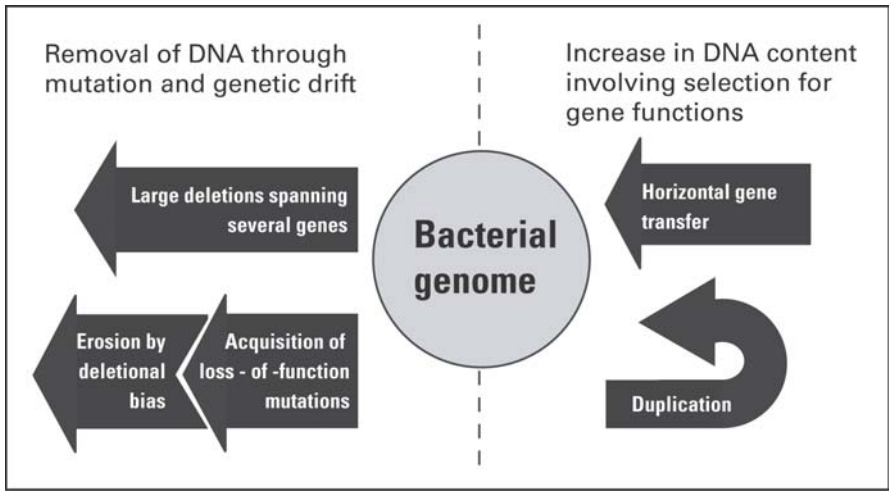



Figure 1. Bacterial genome size is determined by a balance between gene loss and gene gain. A changed life-style might perturb this balance leading to genome reduction. Illustration used with permission from Annika Nilsson. The figure is adapted from Mira et al. (ref 7).

### 1.3 Evolution towards a reduced genome

The range of reported bacterial genome sizes vary from 450 kb to about 10 Mb.[4] The bacteria with the smallest genomes, such as those of *Mycoplasmas*, were previously thought to be more primitive species from which the larger genome species had evolved. Phylogenetic studies have now revealed that the small genome sizes are evolutionary derived from larger genome species through a process of genome reduction.[5] Evolution towards a reduced genome size is associated with a life-style in close contact with a eukaryotic host cell, e.g. endosymbiotic or parasitic bacteria, and not with a specific evolutionary branch. Because bacterial genomes usually contain very little non-coding DNA, over 80 % being open reading frames regardless of genome size, the genome size reduction also results in a reduction in the number of genes encoded.[6,7] It does not seem to be an advantage to have a smaller genome *per se*, in terms of an increased growth rate made possible by a reduced genome replication time.[7] Instead it is the relaxed selection pressure, allowing loss-of-function mutations in previously beneficial genes to form, that lead to genome reduction by deletional bias.

Obligately host-associated bacteria have lost many genes encoding proteins necessary for different metabolic functions. The selection pressure on these genes are relaxed because of the ability of the host cells to supply the bacteria with key metabolites, allowing loss of the components of major biosynthetic pathways, such as genes involved in energy metabolism and synthesis of amino acids, nucleotides and vitamins. The relatively stable environment inside the host also allows for loss of regulatory elements.[8] The extreme genome reduction of the aphid endosymbionts *Buchnera aphidicola* have been extensively studied and it is believed that the gene loss was rapid shortly after the divergence from their free-living ancestors, losing large pieces of DNA in the early stages and that the deletion rate then gradually slowed down.[9] It has been suggested that this large-scale genome reduction might be mediated by recombination at repeated sequences.[10] Host-associated bacteria may

have reduced possibilities to incorporate foreign DNA into their genomes because of their sheltered environment and if so this will also contribute to the genome reduction.[7]

### 1.4 The concept of a minimal genome

What is the minimal number of genes necessary and sufficient to form a living cell? When all essential nutrients are present in unlimited supply and no inhibiting factors such as toxins or competition is present, the majority of the genes in a bacterium, such as *S. typhimurium*, is not essential for sustaining life. A variety of different approaches have been applied for determining such a minimal gene set.

An *in silico* approach of comparing two of the first complete small bacterial genomes sequenced, *Mycoplasma genitalium* and *Haemophilus influenzae*, proved successful. They belong to the gram-positive and gram-negative bacteria respectively, representing evolutionary distant bacterial lineages, which could suggest that genes conserved in both species are essential. Furthermore the genome of *M. genitalium* (580 kb) is one of the smallest known so far and therefore defined an upper bound of the minimal gene set of about 480 genes. The result from this comparison proposed 256 genes as being close to a minimal gene set.[11] As more genomes were completely sequenced another comparative analysis including 21 genomes of bacteria, archaea and one eukaryote were published in the year of 2000. The estimate obtained in this analysis suggested that the minimal gene set might be as small as 150 genes corresponding to an approximate genome size of 150 kb and later comparisons including several endosymbionts and parasites confirmed these results.[8,12] One surprising finding was that a large number of genes seem to be subject to nonorthologous gene displacement, where unrelated or distantly related proteins perform the same function in the cell. A direct comparison of genomes is therefore expected to underestimate the number of genes necessary to sustain cellular life and it is perhaps more fruitful to focus on a minimal set of functions. This means that is does not exist one minimal gene set, but many, each containing all the functions needed for sustaining life.

Several experimental approaches have also been employed for the identification of essential genes under specified growth conditions. These can be divided into strategies focusing on disrupting one gene at the time and those trying to delete large parts of the genome with the goal of experimentally constructing a minimal gene set. The one gene disruption approach to find essential genes have used three different techniques: systematic inactivation of single genes, massive transposon mutagenesis and inhibition of gene expression by antisense RNA.[13,14,15,16,17,18,19] It is important to realize that the essential gene set is not identical with a minimal genome because of redundancy. Several genes can encode for similar functions and will therefore not be simultaneously deletable although they are not defined as essential by single gene inactivation. On the other hand, some genes may not be deletable unless other related genes are deleted simultaneously, a phenomenon that has been observed in large deletions where entire operons are lost. The sizes of the essential gene sets reported using these techniques vary substantially and are larger than those minimal gene sets

obtained by *in silico* approaches. For *Bacillus subtilis* 271 genes were classified as essential and 250 essential genes have been identified in *Escherichia coli*, but the essentiality of many other genes are not determined.[12,20] The essential gene set in *Salmonella* has been estimated to be in the range of 490 genes.[21]

The attempts to engineer bacteria with reduced genomes have so far mainly focused on *Escherichia coli*. No one has yet been able to construct a close to minimal gene set and the largest genome reductions reported have been in the range of a few percent to 29.7 %, corresponding to a maximum loss of 1.38 Mbp.[22,23,24,25] The techniques used to create the deletions include different modified transposons and use of the phage λ Red recombination system.

By combining large amounts of data obtained from the different approaches above, with the notion of a minimal set of functions, a minimal gene set of 206 genes has recently been proposed.[26] This gene set includes almost complete replicational, transcriptional and translational systems, a simple DNA repair system, genes involved in protein and RNA processing, chaperones, a simple transport machinery, genes required for energy generation and biosynthesis of metabolites not present in the environment and structural proteins essential in maintaining cell shape and cell division. The number of genes in each category is shown in Table 2.

Table 2. Functions of minimal gene set genes. Data from Gil *et al.* 2004. [26]

| **Category** | **Subcategory** |
|---|---|
| DNA metabolism (16 genes) | Basic replication machinery (13 genes) |
| | DNA repair, restriction, and modification (3 genes) |
| RNA metabolism (106 genes) | Basic transcription machinery (8 genes) |
| | Translation: aminoacyl-tRNA synthesis (21 genes) |
| | Translation: tRNA maturation and modification (6 genes) |
| | Translation: ribosomal proteins (50 genes) |
| | Translation: ribosome function, maturation and modification (7 genes) |
| | Translation factors (12 genes) |
| | RNA degradation (2 genes) |
| Protein processing, folding, and secretion (15 genes) | Protein posttranslational modification (2 genes) |
| | Protein folding (5 genes) |
| | Protein translocation and secretion (5 genes) |
| | Protein turnover (3 genes) |
| Cellular processes (5 genes) | Cell division (1 gene) |
| | Transport (4 genes) |
| Energetic and intermediary metabolism (56 genes) | Glycolysis (10 genes) |
| | Proton motive force generation (9 genes) |
| | Penthose phosphate pathway (3 genes) |
| | Lipid metabolism (7 genes) |
| | Biosynthesis of nucleotides (15 genes) |
| | Biosynthesis of cofactors (12 genes) |
| Poorly characterized (8 genes) | |
| **Total: 206 genes** | |

Large deletions are possible, in bacteria such as *E. coli* or *S. typhimurium,* because the essential genes are not evenly distributed around the chromosome. The positions of most genes in the minimal gene set proposed by Gil et al. (reference 4) is shown in figure 2. Often, genes encoding related functions are located close together, which can be seen in Figure 2 where many ribosomal proteins are located at 3.6 Mb.

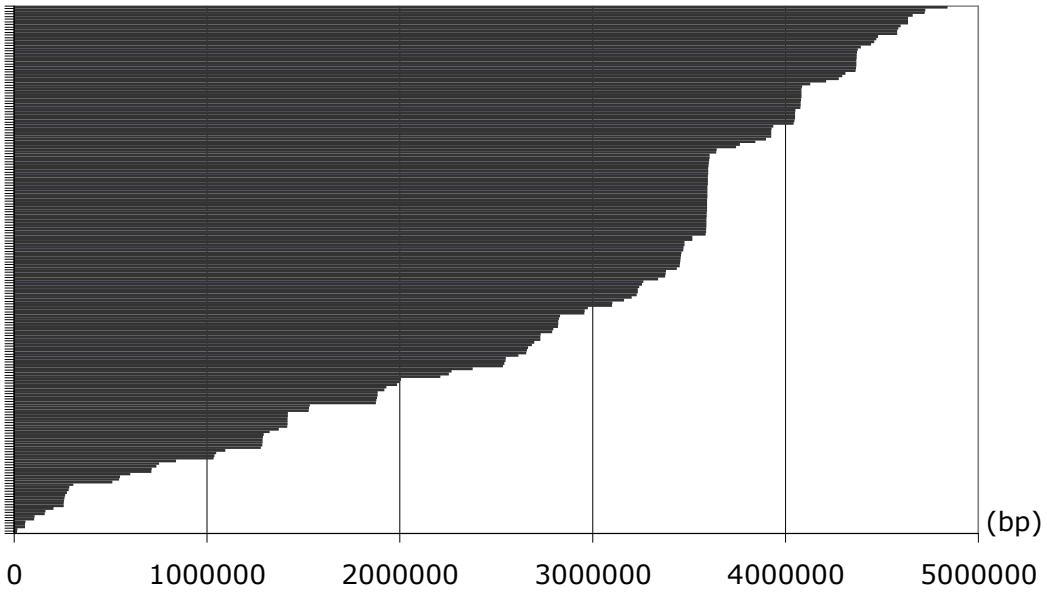



Figure 2. Minimal gene set genes vs. chromosome location (x-axis) in *Salmonella.* Deletions are likely to be possible in areas without minimal gene set genes, shown in the figure as horizontal regions.

**1.5 Experimental evolution**

The short generation time of bacteria allows them to be used to study evolutionary processes in relatively short periods of time. The simplicity of culturing and maintaining bacteria under defined conditions combined with the possibility to freeze them, thereby pausing evolution, make them a good model system for evolutionary studies. The experimentalist is allowed to select the initial genotype as well as the conditions during the experiment, which usually make the results reproducible.

Although one should always be careful in interpreting the results obtained from experimental evolution experiments, because they only demonstrate possible outcomes, it has been shown that it is possible to mimic evolution in nature by experimental evolution.[27] The bacteria species chosen for use in experimental evolution studies should be easily cultured and stored, have a short generation time and be well defined genetically; preferably the complete genome sequence should be available. A variety of genetic tools should also exist to analyze the genetic changes. *S. typhimurium* fulfills all these criteria and were therefore chosen as the species used in this project.

## 1.6 Reduction of genome size in *Salmonella typhimurium*

The aim of this project is to experimentally evolve a *S. typhimurium* bacterium with a reduced genome. A second aim is to study the rate of gene loss and what recombinogenic mechanisms that are involved, which could provide insight into the mechanisms operating in nature. The ultimate goal would be to construct a close to minimal genome, but this cannot be achieved in this short time period, and it is uncertain if the method used here could be used for such a drastic reduction.

A considerably reduced genome can be used to study by which mechanisms a bacterium can gain new gene functions and at what rate. It would also be interesting to study physiological effects of major gene loss, such as changes in the cell size and shape of the bacteria. A bacterium with reduced genome can also be used as a starting point in biotechnology, allowing true genetic engineering by adding only the genes necessary to perform the functions desired. It would also be an advantage to have an extremely well defined genome, encoding less proteins of unknown function, in production of recombinant proteins as well as in the development of a novel, less complex model-organism. Another advantage is that a reduced genome would probably be more stable than the laboratory strains of *E. coli* and *S. typhimurium* commonly used now.

The deleted genes are compared to current knowledge on what genes are essential and what genes are expected to be present in a minimal genome. Large deletable regions are compared with the considerable reduced genomes constructed in *E. coli*, a closely related species. As seen in figure 3 the chromosomal organization of *E. coli* and *S. typhimurium* is very similar, except for an inversion at the replication terminus, so these genomes are easily compared.

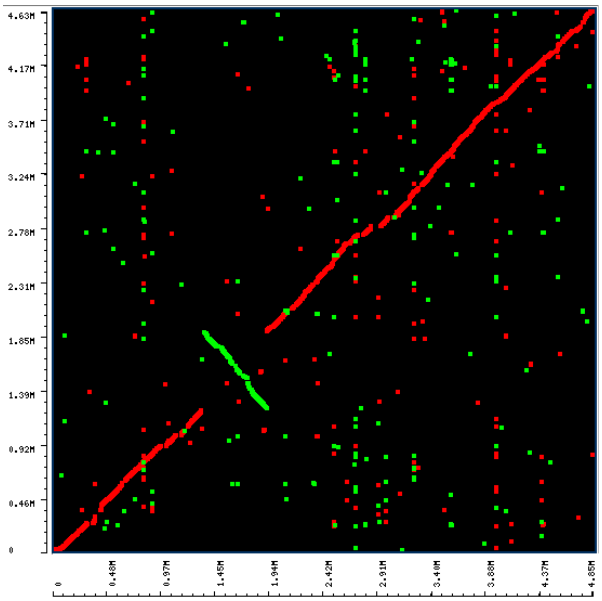



Figure 3: Whole genome alignment of *E. coli* K-12 (y-axis) and *S. typhimurium* (x-axis) using a web-based application of MUMmer.[28] Alignments with the same orientation are shown in red and alignments with opposite orientations are shown in green.

With the method used here, the deletions can be directed to a specific region, which allows rational design of the genome reduction by avoiding regions with known essential genes. A variant of a transposition-defective Tn10, which contain genes for both tetracycline (tetR) and kanamycin resistance (kanR) is used. Mutations leading to loss of tetracycline resistance can be positively selected. Because many of these will be point mutations and small deletions it is also necessary to select for loss of kanamycin resistance. The rationale is that loss of two markers is most likely a result of a large deletion.

Strains with candidate deletions will be investigated using DNA microarrays and pulsed-field gel electrophoresis (PFGE) to confirm the deletions and to find out their sizes and locations. PCR and sequencing will be used to determine the end-points, which allow the recombinogenic mechanism to be studied by looking for the presence of direct repeats indicating RecA-dependent homologous recombination. The absence of homologous sequences on the other hand suggests that the deletions were formed by illegitimate recombination. This can provide insight into what mechanisms are used in nature when larger genome bacteria experience rapid gene loss due to a changed life-style involving close association with eukaryotic cells. The methodology used in the project is schematically shown in Figure 6.

## 2 Materials and methods

### 2.1 Media and antibiotics

Luria-Bertani broth (LB), Luria-Bertani agar plates and antibiotics were prepared and used as previously described.[29] All antibiotics were purchased from Sigma. Bochner plates were used and prepared according to the modifications by Maloy *et al.* of the original protocol.[30,31] Green indicator plates were prepared and used as previously described.[32] All solid media contained 1.5 % agar (BD or Oxoid).

### 2.2 Bacteria strains

The strains used in all experiments are derivatives of *Salmonella enterica* serovar Typhimurium LT2 (referred to in the text as *S. typhimurium*). See Table 1 for details.

### 2.3 Strain construction

Mini transposon Tn10Δ16Δ17 (Tn10dTet) was inserted into the wild-type strain in genes *ydcW* (2 strains with insertions at different locations in the gene), *ydcH*, *srmB*, *uvrB*/*bioD*, *ybgH* and *ychA* by transduction basically as described below. Tn10dTet is a transposition-defective derivative of Tn10, deleted for transposase and left and right IS10. Tn10dTetKan from a F' plasmid, with a kanamycin resistance cassette inserted 148 bp 3' of gene *tetA* in Tn10dTet, was then inserted by transduction to replace the Tn10dTet. (Sequence in Appendix 1)

Transduction with bacteriophage P22 was used to introduce the Tn10 with both tetracycline and kanamycin markers into strains with Tn10dTet at a specific position. The phage lysate prepared by mixing 500 µl overnight culture of the Tn10dTetKan strain with 100 or 500 µl phage P22 ($4*10^6$ pfu/ml, multiplicity of infection 0.004-0.02 phage per cell) and incubating at 37°C overnight with shaking (225 rpm). Cells were then removed by centrifugation (6000 rpm, 10 min) and the supernatant mixed with 500 µl chloroform and centrifuged (13000 rpm, 20 min). The phage lysate was transferred to a new tube and 20 µl chloroform was added to extend the shelf-life of the lysate, which was then stored at 4 °C.

Transduction of strains with a Tn10dTet was performed by mixing 100 µL of overnight cultures with 10 µl of the P22 lysate of the strain with the Tn10dTetKan, followed by phenotypic expression at 37 °C for 90 minutes. The mix was then spread on Luria-Agar (LA) plates containing 30 µg/ml kanamycin and incubated overnight at 37°C to select transductants containing the Tn10dTetKan. Kanamycin resistant colonies were streaked on LA plates with 50 µg/ml kanamycin and on green indicator plates, to be able to select phage-free colonies, and incubated overnight at 37°C. Phage-free colonies were then streaked on green indicator plates through a line of P22 to test that the colonies were not phage-resistant. Colonies confirmed to be phage-free and not phage-resistant were plated on LA plates with 50 µg/ml kanamycin and incubated overnight. Overnight cultures from these plates were stored at –70°C.

To increase the recombination frequency, which also will increase the deletion frequency, two plasmids, each encoding the homologous recombination system of bacteriophage λ, was introduced into strains carrying the Tn10dTetKan. This system, called Red, consists of the proteins encoded by the *exo, bet*, and *gam* genes. These proteins are the λ exonuclease that digests the 5´-ended strand of a dsDNA end, the Beta protein, which binds to ssDNA and mediates strand annealing and the Gam protein that inhibits the activities of the RecBCD and SbcCD nucleases, both involved in bacterial homologous recombination.[33]

The Tn10dTetKan containing bacteria were made electroporation competent by resuspension in 1 ml ice-cold double distilled and autoclaved water ($dH_2O$) followed by centrifugation (6000 rpm, 7 min, 4$^{o}$C); this was repeated three times. The pellets were then resuspended in approximately 100 µl ice-cold $dH_2O$ and 50 µl of the bacteria were mixed with 1 µl of plasmid pSIM5 or pkD46. The mixes were then transferred to cold electroporation cuvettes and electroporated (200 ohm, 25 µF, 2.5 kV). 1ml ice-cold LB broth was added to each cuvette that were kept on ice before transfer to 10 ml tubes that were incubated for 2 h with shaking in 30$^{o}$C. 100 µl of the bacteria was plated on LA plates with chloramphenicol (50 µg/ml) or ampicillin (100 µg/ml) and the rest was centrifuged and the supernatant removed before resuspending in LB broth and plated. Single colonies were restreaked and overnight cultures were grown at 30$^{o}$C from single colonies on the restreaked plates and stored at –70$^{o}$C.

To make sure that the plasmid was present in the strains, plasmid DNA was isolated using QIAprep Spin Miniprep Kit (Qiagen) and analyzed by gel-electrophoresis. The construction and genotypes of all strains are summarized in Table 1, except for the pSIM5 strains constructed analogously to the pKD46 strains.

Table 1: Relevant genotypes of strains used.

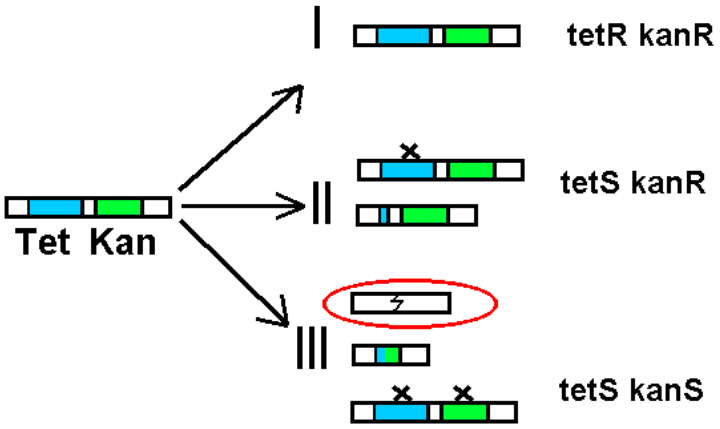
| Strain | Relevant genotype | Recipient | Donor |
|---|---|---|---|
| DA6192 | Wild type | - | - |
| DA6196 | pKD46: araC bla OriR101 repA 101ts (gam+,bet+,exo) | | |
| DA10007 | ydcW::Tn10dTet | DA6192 | DA2309 |
| DA10009 | ydcW::Tn10dTet | DA6192 | DA2315 |
| DA10010 | uvrB/BioD::Tn10dTet | DA6192 | DA3255 |
| DA10014 | ygdH::Tn10dTet | DA6192 | DA3295 |
| DA10015 | ybgH::Tn10dTet | DA6192 | DA3297 |
| DA10018 | srmB::Tn10Ttet | DA6192 | DA3666 |
| DA10020 | ychA::Tn10dTet | DA6192 | DA3677 |
| DA8717 | F'128[pro+,lac+,zzf-1831::Tn10dTetKan]* | | |
| *KanR cassette inserted 148 bp 3' of gene TetA | | | |
| DA10002 | ydcW::Tn10dTetKan | DA10009 | DA8717 |
| DA10003 | ygdH::Tn10dTetKan | DA10014 | DA8717 |
| DA10004 | srmB::Tn10dTetKan | DA10018 | DA8717 |
| DA10005 | ychA::Tn10dTetKan | DA10020 | DA8717 |
| DA10022 | ydcW::Tn10dTetKan | DA10007 | DA8717 |
| | ydcW::Tn10dTetKan + pKD46 | DA10002 | DA6196 |
| | ygdH::Tn10dTetKan + pKD46 | DA10003 | DA6196 |
| | srmB::Tn10dTetKan + pKD46 | DA10004 | DA6196 |
| | ychA::Tn10dTetKan + pKD46 | DA10005 | DA6196 |

## 2.4 Bochner selection

The Bochner selection allows direct, on-plate, positive selection of tetracycline-sensitive (tetS) colonies in a predominantly tetracycline resistant population.[30,31] The inducible tetracycline resistance phenotype (tetR) is mediated by the TetA and TetR proteins; the former is a membrane-bound efflux pump and the latter its repressor. The tetR bacteria are hypersensitive to lipophilic chelators and the inclusion of fusaric acid in the media will inhibit growth of tetR bacteria. The addition of chlortetracycline, which is inactivated by heat denaturation during autoclavation, is necessary to induce expression of the Tet proteins.

Overnight cultures were grown in 96-well microtiter plates at 37$^{o}$C, except for strains with the λ Red system grown at 30$^{o}$C, and 100 μl was spread on Bochner plates. After one or two days colonies appeared. The selection of tetS colonies will pick up all mutations inactivating the Tet proteins. Because only the deletions are of interest here, a screen, where colonies are replica plated on Bochner plates with and without kanamycin (50 μg/ml) is used to increase the likelihood of selecting deletions, because the probability that point-mutations have inactivated both the tetR and kanR genes is small. The replica plated colonies with a tetS and kanS phenotype

were restreaked on Bochner plates with and without kanamycin (50 μg/ml) to reduce the number of false positives (Figure 4). In the initial Bochner selection small colonies were chosen for replica plating, because experiments showed that those were most likely to be deletions. This is reasonable because large deletions are likely to




confer a fitness cost, which leads to a reduced growth rate.[25]

Figure 4: The strategy for the Bochner selection. I. Most bacteria will have functional *tet* genes and will not grow on Bochner plates. II. The replica plating will identify the kanR phenotype of colonies with point-mutations or small deletions in the *tet* genes. III. Only deletions leading to nonfunctional *tet* and *kanR* genes and rare point-mutations in both genes will be investigated further.

Before doing the Bochner selection of the plasmid-containing strains, the λ Red recombination system was induced. The pSIM5 plasmid has a chloramphenicol resistance marker and its λ Red system is under the control of a temperature sensitive promoter; at 30°C no transcripts of the genes are present and at 37°C the expression level is high. Bacteria were cultured at 30°C to OD600≈0.5 in a 96-well microtiter plate in the presence of chloramphenicol and then shifted to 37°C for 15 min. Then a 0.5-2 h temperature shift to 30°C was performed before plating on Bochner plates and grown overnight in 30°C. Colonies were replica plated as described above and incubated overnight at 30°C.

The λ Red system of pkD46 is induced by arabinose, have temperature dependent replication and an ampicillin resistance marker. Overnight cultures (1.5 ml, 50 μg/ml ampicillin, 30°C) were diluted 1:100 in 1 mM L-arabinose LB with ampicillin and grown to OD600≈0.5 (3-4 hours). These cultures were spread on Bochner plates and grown overnight at 37°C without ampicillin, thereby curing the strain of pKD46. All subsequent steps were performed as described above.

### 2.5 Confirmation of Tn10dTetKan deletion by PCR

To confirm the loss of the Tn10dTetKan from the known insertion site the polymerase chain reaction (PCR) was used with primers inside the mini-transposon (for details see Appendix 1). All PCR reactions were performed using GenAmp® PCR System 9700 (Applied Biosystems). At least one positive control was included in all PCR reactions. The PCR template was prepared by suspending one or several

colonies from single-cell streaks in 100 µl $dH_20$ and boiling for five minutes followed by centrifugation (13000 rpm, 5 min).

The PCR mix contained 5 µl of sample template, 0.25 µl (10 mM) of each dNTPs, 10 µl of 10x PCR buffer (Applied Biosystems), 1.6 µl (2mM) $MgCl_2$ (Applied Biosystems), 1 µl (20 pmol) of each primer, 0.1 µl (0.5 U) AmpliTaq Gold polymerase (Applied Biosystems) and 8.3 µl of $dH_2O$ per sample, adding up to a reaction volume of 20 µl. Primers were purchased from MWG-Biotech AG and stored at –20°C. Forward primer sequence: 5´-TCA AGC TCA GGG GAG TAA AC-3´ and reverse primer sequence: 5´-AAA TGG TTG GTC TGC CTT AG-3´. PCR program: initial denaturation for 5 min at 95°C, then 30 cycles of 30s denaturation at 95°C, annealing for 30s at 53-55 °C and elongation for 1 min at 72°C, finished off by extra elongation for 7 min at 72 °C.

PCR-products were analyzed by agarose gel electrophoresis using 1 % SeaKem LE agarose gel (Cambrex) in 1xTBE at 80 V for about 1 hour. The gel was stained with ethidium bromide for 20 min and then examined under ultra-violet light. The positive control was detected as a band at approximately 1.3 kbp consistent with the size of the Tn10dTetKan fragment. The size of the positive control PCR product was determined by comparing with the Generuler® 1 kb ladder (Fermentas).

**2.6 Transduction test**

A preliminary examination of the sizes of the deletions was conducted by transduction with phage P22. Phage lysate was made from the same strain as the deletions were selected from. If the deletion is much smaller than 44kbp, the maximum packaging size of P22, the transduction frequency, as determined by kanR revertants, will be similar to that of the wild-type strain used as control. As shown in Figure 5, larger deletions will have a reduced transduction frequency due to shorter regions of homologous sequences used for recombination.

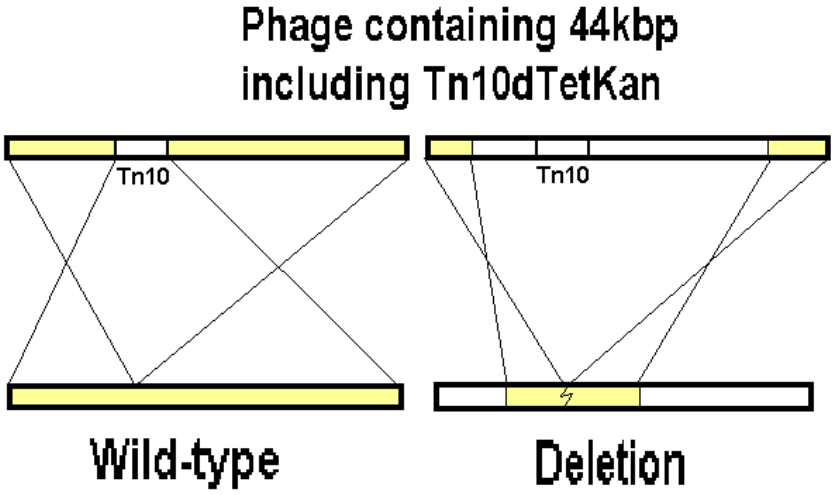



Figure 5: Strains with deletion of the Tn10dTetKan will have shorter or no homologous sequence, leading to a reduction in transduction frequency.

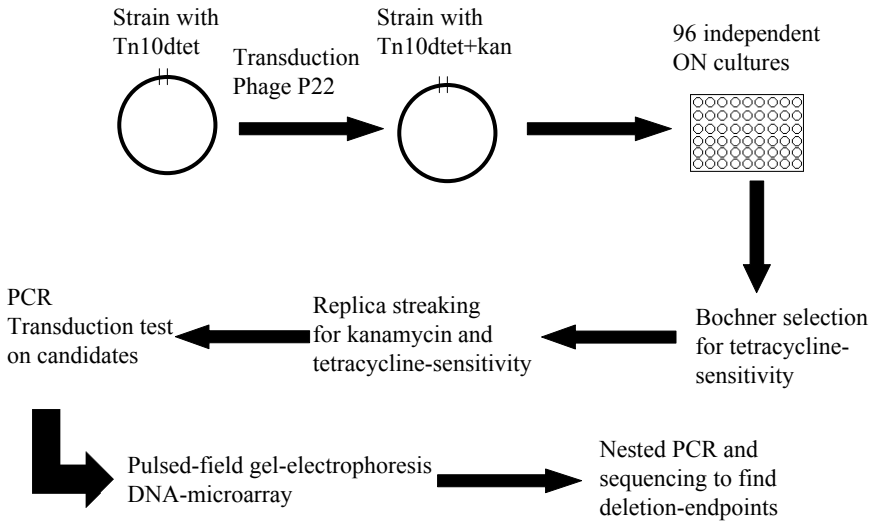



Figure 6. Schematic picture of the methodology used in the project. Mini-transposon were introduced into selected genes of the wild-type strain and then replaced with the mini-transposon carrying two resistance markers. Bochner selection was then used to find tetS strains and subsequently replica plating on Bochner plates with kanamycin allowed the tetS kanS phenotype to be found. PCR confirmed the loss of Tn10dTetKan and the transduction test was used to rule out small deletion before analysis by PFGE and microarray. See text for details.

# 3 Results

## 3.1 Strain construction

*Introduction of Tn10dTetKan into strains with Tn10dTet*

Tn10dTetKan was successfully inserted in five of the seven strains with Tn10dTet, as determined by kanR phenotype and PCR. These five strains, later used for the Bochner selection, had Tn10dTetKan inserted into genes *ydcW, srmB, ygdH* and *ychA*. The strains were confirmed to be phage free and not phage-resistant.

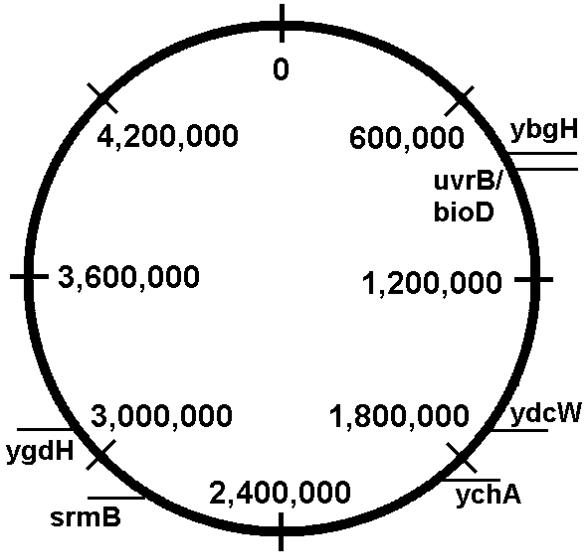



Figure 7. Chromosomal locations of Tn10dTet-insertions.

*Introduction of λ Red System Into Strains With Tn10dTetKan*

The pSIM5 plasmid (contains chloramphenicol resistance marker) with the λ Red system was successfully introduced into strains with Tn10dTetKan inserted into genes *ydcW*, *ygdH* and *ychA*, but not into the strain with insertion in gene srmB. The pKD46 plasmid (contains ampicillin resistance marker) was successfully introduced into strains with Tn10TetKan inserted into genes *ydcW*, *ygdH*, *srmB* and *ychA*.
Plasmid DNA was isolated and the presence of the plasmid, conferring the chloramphenicol or ampicillin resistance phenotype, was confirmed using agarose gel-electrophoresis.

## 3.2 Bochner selection

17 candidate deletions were found using the Bochner selection and replica plating. Six were found in strain DA10002, one in DA10002 with pKD46 and six in strain DA10022; all had Tn10dTetKan inserted in gene *ydcW*. The other four candidate deletions were found in strain DA10004 with the insertion in *srmB*. The number of colonies formed in the initial selection varied substantially between approximately 100 and over 1,000,000 colony forming units (cfu) per ml depending on strain. An

increase in cfu could be observed in strains with the temperature induced λ Red system of pSIM5 grown at 30°C. As this increase could not be fully explained by an increase in deletions, the strains were not used for further Bochner selection experiments.

### 3.3 Confirmation of deletions by PCR

Most candidate deletions that were found to be both tetS and kanS were confirmed to have lost the Tn10dTetKan by PCR with primers inside the mini-transposon. It is expected that the tetS/kanS colonies were in fact deletions as it is unlikely that point-mutations inactivated both genes and false positives from the initial selection are not likely to be kanS.

### 3.4 Transduction tests

In nine strains no transductants were found and this indicates that large deletions were indeed present. Four of these deletions were constructed in DA10004, the strain with Tn10dTetKan inserted into gene *srmB*. Another four were found in DA10022, with insertion in gene *ydcW,* and one was found in the strain with the λ Red plasmid pKD46 and an insertion in *ydcW*. The strains were confirmed not to be phage resistant by using phage lysate of strains with Tn10dTetKan in other positions. Four other strains with a more than 50% reduction in transduction frequency were found, and four strains had between 10 and 50% reduction compared to the wild-type. These strains had Tn10dTetKan inserted in gene *ydcW* and did not harbor the pKD46 plasmid. All strains were confirmed not to be phage resistant.

## 4 Discussion

The replica plating of more than 10000 tetS selected colonies on Bochner plates with and without kanamycin resulted in very few kanamycin sensitive colonies. In earlier experiments, with other strains, about 1% of the tetS colonies were also kanS. There are several possible explanations for this discrepancy. The initial Bochner selection could be non-effective, resulting in a large number of false positives. The bias towards picking small colonies might result in missing most of the deletions. It is unlikely that the large colonies would be large deletions as no deletion was found in large colonies and a large deletion most likely confers a fitness-cost. Perhaps a longer incubation time than two days would allow even smaller colonies to be chosen and that those colonies would contain a larger fraction of deletions. A remote possibility is that the tetracycline resistance genes are not induced quickly enough when plated on Bochner, as the bacteria must be grown in the absence of tetracycline, and that this could result in a larger fraction of tetR colonies. The addition of denatured chlortetracycline, which induces *tetA* gene, but do not affect tetracycline sensitive cells, did not affect the number of cfu indicating that the *tetA* gene are readily induced. By replacing the agar and yeast extract from BD, used for the Bochner plates, with agar and yeast extract from Oxoid, the replica plating was more successful in control experiments with Tn10dTetKan in regions were deletions have previously been found. The selection for tetracycline sensitivity was improved with a reduction of false negatives and this allowed a larger number of deletions to be found by screening a smaller number of colonies. The concentration of kanamycin (50 μg/ml) in the Bochner plates was found not to be high enough to completely inhibit growth as seen when using Luria agar plates. Kanamycin sensitive colonies could be detected because the growth rate of kanS strains were clearly reduced. The presence of Zn ions in the Bochner plates may somehow disturb the antibiotic effect of kanamycin. Another possible explanation for the low frequency of deletion mutants is that the locations of the Tn10dTetKan have been chosen so that deletions are less common in these areas than in earlier strains, and that almost all of the tetS colonies are point-mutations. Strains with Tn10s in other areas should be tested to investigate these problems further.

In the Bochner selection experiments an increase in cfu could be observed in strains with the temperature induced λ Red system of pSIM5. This would be expected as the frequency of recombination is elevated and that would lead to more deletions. The large difference observed here cannot however be satisfactorily explained simply by an increased deletion frequency as control experiments with the recipient strains, without the λ Red system, also exhibit a large increase in cfu when plated at 30$^{o}$C. A more reasonable explanation would be a reduced efficiency of Bochner selection at a lower temperature. The lower growth rate of the λ Red strains was a problem, as they were more difficult to separate from slow growing false positives. Because of these difficulties, the strains carrying the temperature inducible λ Red system encoded on the pSIM5 plasmid, were not used for further Bochner selections. The work was instead focused on the pKD46 carrying strains, which have an arabinose inducible λ Red system.

The long incubation time before slow growing bacteria can be seen on the Bochner plates is a matter of concern as Bochner plates should be used as fresh as possible, preferably within two days. If interesting colonies do not appear within this time period the selection pressure will be reduced and more false positives will appear.

*Locations of deletions and close essential genes*

The sizes of the deletions are yet to be determined by PFGE and microarrays. The deletions obtained were all from the strains with Tn10dTetKan inserted in gene *ydcW*, located at position 1,686,834-1,688,279 bp and gene *srmB* at position 2,782,903-2,784,237. If the *ydcW* region is compared with the corresponding region in *E. coli*, no essential genes are located in the immediate vicinity; the closest are *tyrS*, located at 1.52 Mbp, and *fabI*, at 1.79 Mbp.[20] Large deletions in this area have been constructed in *E. coli*, the largest being a deletion of 300 kbp with the deletion endpoints being very close to the essential genes *tyrS* and *fabI*.[25] This is the largest single deletion found, suggesting that it might be advantageous to continue the search for deletions in this area. Several *S. typhimurium* genes in the area have been reported to be essential.[21] The closest are *narY*, at 1.66 Mbp, and STM1637, at 1.73 Mbp; if these genes are in fact indispensable the largest deletion possible would be less than 60 kbp. All data concerning positions of genes with Tn10dTetKan-insertions and close essential genes are shown in Appendix 2.

The deletions in the area of the *srmB* gene are expected to be larger than 40 kbp because no transductants were found in the transduction test. Two essential *S. typhimurium* genes, *yfiC* and *pssA*, have been reported very close on both sides of *srmB*. If both genes were in fact essential the maximum deletion size would be about 10 kbp. The two closest homologues of essential *E. coli* genes are *lepB* and *pssA,* which are located about 70 kbp apart. Considering that *pssA* are present in both gene sets, as well as in the minimal gene set proposed by Gil *et al.*[26], and that the transduction test indicated a larger than 10 kb deletion, it seems likely that *yfiC* has been deleted. This will of course have to be proven using pulsed-field gel electrophoresis and microarrays. All deletions except one were constructed without use of the λ Red recombination system. This means that they were spontaneous deletions and that analysis of the sequence at the deletion end-points will provide data about which mechanism of recombination that were used.

The positions of the mini-Tn10s were chosen to be just outside regions in which deletions had been made previously. Exceptions are the insertions in genes *ybgH* and *uvrB/bioD,* which are located in an area were deletions have been made earlier.[34] It is unfortunate that the transductions of these strains (to insert the Tn10dTetKan) failed, as they could have been used as positive controls, although the procedure has been validated. It is uncertain if large deletions can be obtained at all the chosen locations. An analysis of reported essential *S. typhimurium* genes and homologues of essential *E. coli* genes reveals that there are closely located essential genes. The possible deletions at these locations are predicted to be less than 15 kbp for the *ychA* gene and less than 40 kbp for *ygdH* (Complete data in Appendix 2). Large deletions made previously have involved several genes classified as essential, so it is possible that the

essential genes located close to the mini-Tn10s also are dispensable as they are not present in the minimal gene set (discussed in section 1.4) proposed by Gil *et al*.[26]

*Future perspectives*

Due to the unexpectedly low number of deletions found, further analysis of the deletions were postponed until more deletion mutants have been constructed. The strains will then be characterized with PFGE to determine the size of the deletion. Whole genome DNA microarrays will be used to determine exactly which genes that have been deleted. This information will be used to make primers for use in semi-random PCR to amplify the DNA sequence over the deletion endpoints. Once this has been achieved the PCR product can be sequenced and this will provide information of the presence of repeats indicating homologous recombination. All these analyses have previously been done for deletions constructed using other techniques, so the analyses are expected to be straightforward. More strains should be constructed with Tn10dTetKan at other locations, including known deletable regions.

## 5 Acknowledgements

I would like to thank Professor Dan Andersson for giving me the opportunity to work in his group and for interesting scientific discussions. I am grateful to Sanna Koskiniemi for the help and guidance with both experimental work and preparation of this thesis. I would also like to thank Linus Sandegren, Eric Wiktelius and Andreas Cederlund for critical reading of the manuscript. Thanks also to the rest of the Dan Andersson group and everyone at UGSBR.

# 6 Appendices

## Appendix 1

Sequence of Tn10dTetKan.

Tn10dTet in capital letters and *kanR* gene, inserted at base 2961 of Tn10dTet, in lower-case letters. The positions for primers used are marked in grey. The bold letters shows the beginning and end of *tetR* and *tetA* genes.

CTGATGAATCCCCTAATGATTTTTATCAAAATCATTAAGTTAAGGTAGATACACATCTTGTCATATGATCT
ATGATTCCCTTTGTCAACAGCAATGGATCACTGAAAATGGTTCAATGATCACATTAAGTGGTATTCAATA
TTTTCATGAAATGGGAATTGACGTTCCTTCCAAACATTCACGTAAAATCTGTTGTGCGTGTTTAGATTGGA
GTGAACGCCGTTTCCATTTAGGTGGGTACGTTGGAGCCGCATTATTTTCGCTTTATGAATCTAAAGGGTGG
TTAACTCGACATCTTGGTTACCGTGAAGTTACCATCACGGAAAAAGGTTATGCTGCTT**T**TAAGACCCACT
TTCACATTTAAGTTGTTTTTCTAATCCGCATATGATCAATTCAAGGCCGAATAAGAAGGCTGGCTCTGCAC
CTTGGTGATCAAATAATTCGATAGCTTGTCGTAATAATGGCGGCATACTATCAGTAGTAGGTGTTTCCCTT
TCTTCTTTAGCGACTTGATGCTCTTGATCTTCCAATACGCAACCTAAAGTAAAATGCCCCACAGCGCTGAG
TGCATATAATGCATTCTCTAGTGAAAAACCTTGTTGGCATAAAAAGGCTAATTGATTTTCGAGAGTTTCAT
ACTGTTTTTCTGTAGGCCGTGTACCTAAATGTACTTTTGCTCCATCGCGATGACTTAGTAAAGCACATCTA
AAACTTTTAGCGTTATTACGTAAAAAATCTTGCCAGCTTTCCCCTTCTAAAGGGCAAAAGTGAGTATGGT
GCCTATCTAACATCTCAATGGCTAAGGCGTCGAGCAAAGCCCGCTTATTTTTTACATGCCAATACAATGT
AGGCTGCTCTACACCTAGCTTCTGGGCGAGTTTACGGGTTGTTAAACCTTCGATTCCGACCTCATTAAGCA
GCTCTAATGCGCTGTTAATCACTTTACTTTTATCTAATCTAGACATCATTAATTCCTAATTTTTG**T**TGACAC
TCTATCATTGATAGAGTTATTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAA**A**TGAATAGTTCGACA
AAGATCGCATTGGTAATTACGTTACTCGATGCCATGGGGATTGGCCTTATCATGCCAGTCTTGCCAACGTT
ATTACGTGAATTTATTGCTTCGGAAGATATCGCTAACCACTTTGGCGTATTGCTTGCACTTTATGCGTTAA
TGCAGGTTATCTTTGCTCCTTGGCTTGGAAAAATGTCTGACCGATTTGGTCGGCGCCCAGTGCTGTTGTTG
TCATTAATAGGCGCATCGCTGGATTAGTTATTGCTGGCTTTTTCAAGTGCGCTTTGGATGCTGTATTTAGG
CCGTTTGCTTTCAGGGATCACAGGAGCTACTGGGGCTGTCGCGGCATCGGTCATTGCCGATACCACCTCA
GCTTCTCAACGCGTGAAGTGGTTCGGTTGGTTAGGGGCAAGTTTTGGGCTTGGTTTAATAGCGGGGCCTA
TTATTGGTGGTTTTGCAGGAGAGATTTCACCGCATAGTCCCTTTTTTATCGCTGCGTTGCTAAATATTGTC
ACTTTCCTTGTGGTTATGTTTTGGTTCCGTGAAACCAAAAATACACGTGATAATACAGATACCGAAGTAG
GGGTTGAGACGCAATCGAATTCGGTATACATCACTTTATTTAAAACGATGCCCATTTTGTTGATTATTTAT
TTTTCAGCGCAATTGATAGGCCAAATTCCCGCAACGGTGTGGGTGCTATTTACCGAAAATCGTTTTGGAT
GGAATAGCATGATGGTTGGCTTTTCATTAGCGGGTCTTGGTCTTTTACACTCAGTATTCCAAGCCTTTGTG
GCAGGAAGAATAGCCACTAAATGGGGCGAAAAAACGGCAGTACTGCTCGAATTTATTGCAGATAGTAGT
GCATTTGCCTTTTTAGCGTTTATATCTGAAGGTTGGTTAGATTTCCCTGTTTTAATTTTATTGGCTGGTGGT
GGGATCGCTTTACCTGCATTACAGGGAGTGATGTCTATCCAAACAAAGAGTCATGAGCAAGGTGCTTTAC
AGGGATTATTGGTGAGCCTTACCAATGCAACCGGTGTTATTGGCCCATTACTGTTTACTGTTATTTATAAT
CATTCACTACCAATTTGGGATGGCTGGATTTGGATTATTGGTTTAGCGTTTTACTGTATTATTATCCTGCTA
TCGATGACCTTCATGTTAACCCCTCAAGCTCAGGGGAGTAAACAGGAGACAAGTGCTTA**G**TTATTTCGTC
ACCAAATGATGTTATTCCGCGAAATATAATGACCCTCTTGATAACCCAAGAGGGCATTTTTTACGATAAA
GAAGATTTAGCTTCAAATAAAACCTATCTATTTTATTTATCTTTCAAGCTCAATAAAAAGCCGCGGTAtccg
gccgcttgggtggagaggctattcggctatgactgggcacaacagacaatcggctgctctgatgccgccgtgttccggctgtcagcgcaggggcgcccggttcttttttgtcaa
gaccgacctgtccggtgccctgaatgaactgcaggacgaggcagcgcggctatcgtggctggccacgacgggcgttccttgcgcagctgtgctcgacgttgtcactgaagc
gggaagggactggctgctattgggcgaagtgccggggcaggatctcctgtcatctcaccttgctcctgccgagaaagtatccatcatggctgatgcaatgcggcggctgcat
acgcttgatccggctacctgcccattcgaccaccaagcgaaacatcgcatcgagcgagcacgtactcggatggaagccggtcttgtcgatcaggatgatctggacgaagag
catcaggggctcgcgccagccgaactgttcgccaggctcaaggcgcgcatgcccgacggcgaggatctcgtcgtgacccatggcgatgcctgcttgccgaatatcatggtg
gaaaatggccgcttttctggattcatcgactgtggccggctgggtgtggcggaccgctatcaggacatagcgttggctacccgtgatattgctgaagagcttggcggcgaatg
ggctgaccgcttcctcgtgctttacggtatcgccgctcccgattcgcagcgcatcgccttctatcgccttcttgacgagttcttctgagcgggacAATAGCAATAAA
TTGGCCTTTTTTATCGGCAAGCTCTTTTAGGTTTTTCGCATGTATTGCGATATGCATAAACCAGCCATTGA
GTAAGTTTTTAAGCACATCATCATCATAAGCTTTAAGTTGGTTCTCTTGGATCAATTTGCTGACAATGGCG
TTTACCTTACCAGTAATGTATTCAAGGCTAATTTTTTCAAGTTCATTCCAACCAATGATAGGCATCACTTC
TTGGATAGGGATAAGGTTTTTATTATTATCAATAATATAATCAAGATAATGTTCAAATATACTTTCTAAGG
CAGACCAACCATTTGTTAAATCAGTTTTTGTTGTGATGTAGGCATCAATCATAATTAATTGCTGCTTATAA
CAGGCACTGAGTAATTGTTTTTTATTTTTAAAGTGATGATAAAAGGCACCTTTGGTCACCAACGCTTTTCC
CGAGATCATATGACAAGATGTGTATCCACCTTAACTTAATGATTTTTACCAAAATCATTAGGGGATTCATC
AG

Appendix 2. Locations of genes chosen for deletion experiments (in bold) and close essential genes

| Gene number | Gene symbol | Location begin. | Location end |
|---|---|---|---|
| **Genes with Tn10dTetKan insertion and essential genes in *S. typhimurium*. (ref 21, Knuth *et al.*)** | | | |
| STM1578 | *narY* | 1664425 | 1665969 |
| **STM1597** | ***ydcW*** | **1686834** | **1688279** |
| STM1637 | | 1726876 | 1728426 |
| STM1763 | *narH* | 1856881 | 1858416 |
| **STM1773** | ***ychA*** | **1871299** | **1872108** |
| STM1775 | *hemK* | 1872498 | 1873331 |
| STM2642 | *yfiC* | 2782035 | 2782772 |
| **STM2643** | ***srmB*** | **2782903** | **2784237** |
| STM2652 | *pssA* | 2792655 | 2794010 |
| STM2962 | *gudT* | 3113237 | 3114595 |
| **STM2969** | ***ygdH*** | **3118842** | **3120206** |
| STM2996 | *recC* | 3152712 | 3156083 |
| **Genes with Tn10dTetKan insertion and homologues of essential *E. coli* genes. (ref. 20, PEC)** | | | |
| STM1449 | *tyrS* | 1523487 | 1524761 |
| **STM1597** | ***ydcW*** | **1686834** | **1688279** |
| STM1700 | *fabI* | 1792136 | 1792924 |
| STM1772 | *kdsA* | 1870407 | 1871261 |
| **STM1773** | ***ychA*** | **1871299** | **1872108** |
| STM1776 | *prfA* | 1873331 | 1874413 |
| STM2582 | *lepB* | 2725781 | 2726755 |
| **STM2643** | ***srmB*** | **2782903** | **2784237** |
| STM2652 | *pssA* | 2792655 | 2794010 |
| STM2952 | *eno* | 3097364 | 3098662 |
| **STM2969** | ***ygdH*** | **3118842** | **3120206** |
| STM3002 | *lgt* | 3158818 | 3159693 |
| **Genes with Tn10dTet insertion and essential genes in *S. typhimurium*. (ref. 21, Knuth *et al.*)** | | | |
| STM0686 | *glnS* | 747421 | 749088 |
| **STM0710** | ***ybgH*** | **776238** | **777719** |
| STM0720 | | 785165 | 786061 |
| STM0791 | *hutH* | 856655 | 858175 |
| **STM0798** | ***uvrB*** | **864433** | **866454** |
| STM0813 | *ybhP* | 878350 | 879108 |
| **Genes with Tn10dTet insertion and homologues of essential *E. coli* genes. (ref. 20, PEC)** | | | |
| STM0694 | *fldA* | 757128 | 757658 |
| **STM0710** | ***ybgH*** | **776238** | **777719** |
| STM0953 | *infA* | 1030347 | 1030565 |
| STM0694 | *fldA* | 757128 | 757658 |
| **STM0798** | ***uvrB*** | **864433** | **866454** |
| STM0953 | *infA* | 1030347 | 1030565 |

# 7 References


[1] *Salmonella* information. salmonella.org/info.html, 8 May 2005.

[2] Pang T, Levine MM, Ivanoff B, Wain J, Finlay BB. 1998. Typhoid fever--important issues still remain. Trends Microbiol. 6(4):131-3.

[3] McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, Hou S, Layman D, Leonard S, Nguyen C, Scott K, Holmes A, Grewal N, Mulvaney E, Ryan E, Sun H, Florea L, Miller W, Stoneking T, Nhan M, Waterston R, Wilson RK. 2001. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. Nature 413(6858):852-6.

[4] Gil R, Sabater-Munoz B, Latorre A, Silva FJ, Moya A. 2002. Extreme genome reduction in *Buchnera spp*.: toward the minimal genome needed for symbiotic life. Proc Natl Acad Sci U S A. 99(7):4454-8.

[5] Woese CR. 1987. Bacterial Evolution. Microbiol Rev. 51(2):221-71.

[6] Konstantinidis KT, Tiedje JM. 2004. Trends between gene content and genome size in prokaryotic species with larger genomes. Proc Natl Acad Sci U S A. 101(9):3160-5.

[7] Mira A, Ochman H, Moran NA. 2001. Deletional bias and the evolution of bacterial genomes. Trends Genet. 17(10):589-96.

[8] Moran NA. 2002. Microbial minimalism: genome reduction in bacterial pathogens. Cell 108(5):583-6.

[9] Tamas I, Klasson L, Canback B, Naslund AK, Eriksson AS, Wernegreen JJ, Sandstrom JP, Moran NA, Andersson SG. 2002. 50 million years of genomic stasis in endosymbiotic bacteria. Science 296(5577):2376-9.

[10] Klasson L, Andersson SG. 2004. Evolution of minimal-gene-sets in host-dependent bacteria. Trends Microbiol. 12(1):37-43.

[11] Mushegian AR, Koonin EV. 1996. A minimal gene set for cellular life derived by comparison of complete bacterial genome. Proc Natl Acad Sci U S A. 93(19):10268-73.

[12] Koonin, EV. 2000. How many genes can make a cell: the minimal-gene-set concept.
Annu Rev Genomics Hum Genet. 1:99-116.

[13] Ji Y, Zhang B, Van SF, Horn, Warren P, Woodnutt G, Burnham MK, Rosenberg M. 2001. Identification of critical staphylococcal genes using conditional phenotypes generated by antisense RNA. Science 21;293(5538):2266-9.

[14] Forsyth RA, Haselbeck RJ, Ohlsen KL, Yamamoto RT, Xu H, Trawick JD, Wall D, Wang L, Brown-Driver V, Froelich JM, C KG, King P, McCarthy M, Malone C, Misiner B, Robbins D, Tan Z, Zhu Zy ZY, Carr G, Mosca DA, Zamudio C, Foulkes JG, Zyskind JW. 2002. A genome-wide strategy for the identification of essential genes in Staphylococcus aureus. Mol Microbiol. 43(6):1387-400.

[15] Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, Arkin AP, Astromoff A, El-Bakkoury M, Bangham R, Benito R, Brachat S, Campanaro S, Curtiss M, Davis K, Deutschbauer A, Entian KD, Flaherty P, Foury F, Garfinkel DJ, Gerstein M, Gotte D, Guldener U, Hegemann JH, Hempel S, Herman Z, Jaramillo DF, Kelly DE, Kelly SL, Kotter P, LaBonte D, Lamb DC, Lan N, Liang H, Liao H, Liu L, Luo C, Lussier M, Mao R, Menard P, Ooi SL, Revuelta JL, Roberts CJ, Rose M, Ross-Macdonald P, Scherens B, Schimmack G, Shafer B, Shoemaker DD, Sookhai-Mahadeo S, Storms RK, Strathern JN, Valle G, Voet M, Volckaert G, Wang CY, Ward TR, Wilhelmy J, Winzeler EA, Yang Y, Yen G, Youngman E, Yu K, Bussey H, Boeke JD, Snyder M, Philippsen P, Davis RW, Johnston M. 2002. Functional profiling of the *Saccharomyces cerevisiae* genome. Nature 25;418(6896):387-91.

[16] Kobayashi K, Ehrlich SD, Albertini A, Amati G, Andersen KK, Arnaud M, Asai K, Ashikaga S, Aymerich S, Bessieres P, Boland F, Brignell SC, Bron S, Bunai K, Chapuis J, Christiansen LC, Danchin A, Debarbouille M, Dervyn E, Deuerling E, Devine K, Devine SK, Dreesen O, Errington J, Fillinger S, Foster SJ, Fujita Y, Galizzi A, Gardan R, Eschevins C, Fukushima T, Haga K, Harwood CR, Hecker M, Hosoya D, Hullo MF, Kakeshita H, Karamata D, Kasahara Y, Kawamura F, Koga K,

Koski P, Kuwana R, Imamura D, Ishimaru M, Ishikawa S, Ishio I, Le Coq D, Masson A, Mauel C, Meima R, Mellado RP, Moir A, Moriya S, Nagakawa E, Nanamiya H, Nakai S, Nygaard P, Ogura M, Ohanan T, O'Reilly M, O'Rourke M, Pragai Z, Pooley HM, Rapoport G, Rawlins JP, Rivas LA, Rivolta C, Sadaie A, Sadaie Y, Sarvas M, Sato T, Saxild HH, Scanlan E, Schumann W, Seegers JF, Sekiguchi J, Sekowska A, Seror SJ, Simon M, Stragier P, Studer R, Takamatsu H, Tanaka T, Takeuchi M, Thomaides HB, Vagner V, van Dijl JM, Watabe K, Wipat A, Yamamoto H, Yamamoto M, Yamamoto Y, Yamane K, Yata K, Yoshida K, Yoshikawa H, Zuber U, Ogasawara N. 2003. Essential *Bacillus subtilis* genes. Proc Natl Acad Sci U S A. 100(8):4678-83.

[17] Herring CD, Glasner JD, Blattner FR. 2003. Gene replacement without selection: regulated suppression of amber mutations in *Escherichia coli*. Gene 311:153-63.

[18] Mori H, Isono K, Horiuchi T, Miki T. 2000. Functional genomics of *Escherichia coli* in Japan. Res Microbiol. 151(2):121-8.

[19] Hutchison CA, Peterson SN, Gill SR, Cline RT, White O, Fraser CM, Smith HO, Venter JC. 1999. Global transposon mutagenesis and a minimal *Mycoplasma* genome. Science 286(5447):2165-9.

[20] PEC (Profiling of *E. coli* chromosome) http://www.shigen.nig.ac.jp/ecoli/pec/index.jsp. (8 May 2005.

[21] Knuth K, Niesalla H, Hueck CJ, Fuchs TM. Large-scale identification of essential *Salmonella* genes by trapping of lethal insertions. Molecular Microbiology. 2004 51(6) 1729-44

[22] Kolisnychenko V, Plunkett G 3rd, Herring CD, Feher T, Posfai J, Blattner FR, Posfai G. 2002. Engineering a reduced *Escherichia coli* genome. Genome Res. 12(4):640-7.

[23] Yu BJ, Sung BH, Koob MD, Lee CH, Lee JH, Lee WS, Kim MS, Kim SC. 2002. Minimization of the *Escherichia coli* genome using a Tn5-targeted Cre/loxP excision system. Nat Biotechnol. 20(10):1018-23.

[24] Goryshin IY, Naumann TA, Apodaca J, Reznikoff WS. 2003. Chromosomal deletion formation system based on Tn5 double transposition: use for making minimal genomes and essential gene analysis. Genome Res. 13(4):644-53.

[25] Hashimoto M, Ichimura T, Mizoguchi H, Tanaka K, Fujimitsu K, Keyamura K, Ote T, Yamakawa T, Yamazaki Y, Mori H, Katayama T, Kato J. 2005. Cell size and nucleoid organization of engineered *Escherichia coli* cells with a reduced genome. Mol Microbiol. 55(1):137-49.

[26] Gil R, Silva FJ, Pereto J, Moya A. 2004. Determination of the core of a minimal bacterial gene set. Microbiol Mol Biol Rev. 68(3):518-37.

[27] Souza V, Travisano M, Turner PE, Eguiarte LE. 2002. Does experimental evolution reflect patterns in natural populations? *E. coli* strains from long-term studies compared with wild isolates. Antonie Van Leeuwenhoek. 81(1-4):143-53.

[28] The Institute for Genomic Research. http://www.tigr.org/tigr-scripts/CMR2/webmum/mumplot. 12 May 2005. Fully described in A.L. Delcher, S. Kasif, R.D. Fleischmann, J. Peterson, O. White, and S.L. Salzberg. 1999. Alignment of whole genomes. Nucleic Acids Research, 27:11, 2369-2376.

[29] Davis, R. W., Botstein D., Roth J. R., Advanced Bacterial Genetics. 1980. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[30] Maloy SR, Nunn WD. Selection for loss of tetracycline resistance by *Escherichia coli*. 1981. J Bacteriol. 145(2):1110-1.

[31] Bochner BR, Huang HC, Schieven GL, Ames BN. 1980. Positive selection for loss of tetracycline resistance. J Bacteriol. 1980 Aug;143(2):926-33.

[32] Davis, R. W., D. Botstein, and J. R. Roth. 1980. Advanced bacterial genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

[33] Poteete AR. 2001. What makes the bacteriophage lambda Red system useful for genetic engineering: molecular mechanism and biological function. FEMS Microbiol Lett. 201(1):9-14.

[34] Personal communication. Sanna Koskiniemi. Department of Medical Biochemistry and Microbiology. Uppsala University.