CHAPTER 6

UPGRADING THE TOOLKIT OF LACTOCOCCUS LACTIS§

§João P. C. Pinto, Araz Zeyniyev, Sjoerd van der Meulen, Oliwia Bieniek, Harma Karsens, Hein Trip, Juke S. Lolkema, Oscar P. Kuipers and Jan Kok

Part of this chapter was published in:

Abstract

*Lactococcus lactis* is an industrially important microorganism and a model for the family of Lactic Acid Bacteria (LAB). Catering to the ever increasing need to constantly update the genetic toolkit of *L. lactis*, we constructed plasmid pSEUDO and derivatives and used them to show that *llmg_pseudo_10* in *Lactococcus lactis* MG1363, and its homologous locus in *L. lactis* IL1403, are suitable for chromosomal integrations. *Lacticis* MG1363 and IL1403 NICE-system-derivatives (JP9000 and IL9000) and two general stress reporter-strains (NZ9000::P*hrcA*-GFP; NZ9000::PgrosGFP) enabling *in-vivo* non-invasive monitoring of cellular fitness were constructed. We additionally established a protocol to easily screen and characterize randomly produced mutants of this organism. We successfully implemented and validated a protocol for Genomic Array Footprinting (GAF), using SuperAmine glass slides spotted with duplicates of around 2500 ORF amplicons of *L. lactis* subsp. *cremoris* MG1363. We additionally constructed *pGh8T7:IS1*, a plasmid with a tetracycline resistance marker rather than the often used erythromycin resistance that is present in *pGh9T7:IS1*. This unlocks experimental setup and screening strategies on strains that are already erythromycin resistant, *e.g.* from carrying a genetic construct used in the selection procedure.


**Introduction**

*Lactococcus lactis* is an industrially important microorganism which is widely used in dairy fermentation. Glancing into the future, this bacterium could also play a role in other, novel, biotechnological applications, for instance as a possible membrane protein overproduction host, as is explored in the work described in this thesis.

*L. lactis* has also been a model organism for the Lactic Acid Bacteria (LAB) for the past several decades and the current knowledge regarding the molecular biology of this family of bacteria derives in great part from evidence initially collected from *L. lactis*. With the public availability of eight genome sequences of this bacterium (1–8), *L. lactis* also leads the way with respect to post-genomics experimental approaches such as transcriptomics, proteomics and metabolomics and using these for in-depth systems biology studies (see for example (9)).

The ever increasing pace of scientific research on the LAB demands continuously updating the genetics and genomics toolkit of *L. lactis* so that it remains a valid model organism in which experimental questions can be readily addressed and answered. Here we document the construction a standard for chromosomal integrations and the set up a protocol in *L. lactis* for the easy screening and characterization of randomly produced mutants.

A standard locus in the chromosome of *L. lactis* for integration of DNA fragments, whether for genetic complementation (single-copy or merodiploid-like situations) or cloning of a reporter gene or -promoter, is lacking. Contrary to *e.g.* *Bacillus subtilis*, where the *amyE* locus is often used for these purposes (22), in *L. lactis* various chromosomal loci have been chosen as targets for integration. Unfortunately, and to the best of our knowledge, none of the proposed strategies exclude the possibility of phenotypic consequences. The *leuA* locus was shown to suffer from active read-through from the native *leuA* promoter (11), whereas the choice of the sex factor locus might interfere with the biology of *L. lactis* and, in addition, may have consequences with respect to possible conjugational transfer of the inserted DNA. Also, bacteriophage sequences have been used to drive site-specific integration of plasmids in the chromosome of *L. lactis* (3, 9). However, these methods do not allow making strains without resistance markers while some require the use of a second plasmid to provide the bacteriophage integrase *in trans*. Furthermore, the localization in the chromosome of *L. lactis* MG1363 of sequences with high similarity to a given *attB*, e.g. in the *comGC* gene for TP901-1 *attB* and in *rex* for TUC2009 *attB* (data not shown), implies that the integration process might lead to the simultaneous disruption of potentially relevant processes.

Here it was examined whether the *llmg_pseudo_10* locus of *L. lactis* MG1363, or its corresponding region in *L. lactis* IL1403, is a suitable neutral region for chromosomal
integrations (Figure 4). In *L. lactis* IL1403, this locus contains *yfjF*, a gene of 1506 bp of which the product exhibits homology to transport proteins from the major facilitator superfamily. In *L. lactis* MG1363, translation is halted prematurely due to the presence of a stop codon at position 303, hence its annotation as a pseudo gene (24). By cloning and re-sequencing, the nucleotide sequence of this region in *L. lactis* MG1363, originally described by Wegmann et al. (24), was confirmed. The loss of function of the locus in *L. lactis* MG1363 suggests that *yfjF* is non-essential. The *llmg_pseudo_10* locus has been shown to be silent throughout the growth of *L. lactis* MG1363 in batch cultures in M17 medium (4) and milk (Anne de Jong, personal communication). In addition, *llmg_pseudo_10* and *yfjF* display low nucleotide sequence similarity with other regions in the *L. lactis* genome, minimizing the possibility of illegitimate recombination.

![Genomic context of *llmg_pseudo_10* of *L. lactis* MG1363 and its relation to the homologous regions (hr’s) present in pSEUDO and pSEUDO-GFP.](image)

**Figure 1** – Genomic context of *llmg_pseudo_10* of *L. lactis* MG1363 and its relation to the homologous regions (hr’s) present in pSEUDO and pSEUDO-GFP. The multiple cloning site (MCS) contains, in clockwise order, EcoRI, Xmal/Smal, SpII, ScaII, SalI, HindIII, BgIII, XhoI and BamHI restriction enzyme sites. The *gfp* gene for super-folder GFP (19) was cloned in pSEUDO using the XhoI and BamHI sites, yielding pSEUDO-GFP. Terminators are indicated by ‘lollipop’ structures. The vertical line on *llmg_pseudo_10* depicts the stop codon that prematurely halts translation of the gene in *L. lactis* MG1363. *eryR*: erythromycin resistance gene; *oroP* encodes the *L. lactis* orotate transporter (5).
Part I: pSEUDO, a genetic integration standard for Lactococcus lactis

Construction and applicability of pSEUDO and pSEUDO-GFP

To be able to perform unmarked integrations in the chromosome of *L. lactis* the chromosomal integration vector pCS1966 was employed, allowing to positively select cells in which the plasmid has been excised from the genome (16). Two DNA fragments were amplified from *L. lactis* MG1363 chromosomal DNA by PCR, one of 529 bp obtained with the primer pair P1_pseudo10/P2_pseudo10 and another of 804 bp using P3_pseudo10/P4_pseudo10, and sequentially inserted into pCS1966 using the restriction enzymes indicated in Table 1, and *E. coli* DH5α as the cloning host. Selection was performed on TY agar plates with 150 µg/ml erythromycin. The custom-made multiple cloning site GAATTCCCCGGGTCGACGCTTCGGGTCGAGGATCC was introduced between the BamHI and EcoRI sites, thus producing the plasmid pSEUDO (Figure 4). This vector can be used to insert DNA fragments into the *llmg_pseudo_10* locus of *L. lactis* MG1363 using positive selection for resistance to the toxic pyrimidine analog 5-fluoroorotate, as described before (16), with minor modifications (17). Although it is not possible to screen for integrants via loss of function through gene disruption, pSEUDO allows for a quick and efficient positive survival strategy to monitor both integration (erythromycin resistance) and excision of the vector backbone from the chromosome (5-fluoroorotate resistance), thus enabling the production of unmarked strains in an easy and fast manner.

As a proof of principle, the applicability of pSEUDO is illustrated through the integration of the genes of the two-component nisin sensor, NisRK (12), in the chromosome of *L. lactis* MG1363, internal and in opposite orientation to *llmg_pseudo_10*. A fragment containing the nisRK genes with their own promoter was amplified by PCR from chromosomal DNA of *L. lactis* NZ9000 (20) using the primers nisRK_Forw and nisRK_Rev. The PCR product and pSEUDO were digested with BamHI and ligated after dephosphorylation of the digested vector, producing pSEUDO::nisRK. The presence and orientation of the insert were checked with PCR and restriction endonuclease digestions. The integration of nisRK into *L. lactis* MG1363 using pSEUDO::nisRK generated the *L. lactis* strain JP9000. Expression of nisRK and the applicability of the nisin-induced controlled expression (NICE) system (18, 19) using *L. lactis* JP9000 are the same as for *L. lactis* NZ9000. As an example, the NICE system was used for the overproduction of a GFP- and hexa-histidine-tagged membrane protein of *L. lactis*, BcaP-GFP-H6, using pNZbcaP-GFP-H6, a pNZ8048 derivative in which the bcaP-gfp-H6 gene is driven by the nisin-inducible PnisA promoter (11). *L. lactis* strains NZ9000 and JP9000 carrying this plasmid produced equivalent levels of the tagged protein since a similar fluorescent signal from the overproduced BcaP-GFP-H6 was obtained using either strain
Chapter 6

(Figure 2). Contrary to previous observations (20), the native nisRK promoter (21) was sufficient to yield significant amounts of NisRK and read-through from the neighboring genes was not required for a functional NICE system. In addition to pSEUDO::nisRK, a similar vector was constructed for nisRK integration in the yffF locus of *L. lactis* IL1403. The *llmg_pseudo_10* homologous regions of pSEUDO were replaced by homologous regions of the *yffF* locus, amplified from chromosomal DNA of IL1403 using primers P1_yffF, P2_yffF, P3_yffF and P4_yffF (Table 1). The nisRK genes were inserted in the BamHI site in opposite direction to the *yffF* gene, resulting in plasmid pSEUDO-IL::nisRK. The integration of nisRK into *L. lactis* IL1403 resulted in strain IL9000. To test functionality, IL9000 was transformed with a pNZ8048-derived plasmid in which *gfp* was inserted downstream the nisin-inducible P*nisA* promoter (H. Trip, unpublished results). GFP fluorescence levels upon induction with nisin were similar to those obtained with *L. lactis* NZ9000 harboring the same plasmid (data not shown), demonstrating that the NICE system is identically functional in *L. lactis* IL9000. The plasmids pSEUDO::nisRK and pSEUDO-IL::nisRK can be used to introduce the NICE system in any *L. lactis* MG1363- or *L. lactis* IL1403-derived strain, respectively. This is thought to be very useful for the LAB research community since *L. lactis* MG1363 and *L. lactis* IL1403 are by far the most commonly used strains in the *L. lactis* applied and fundamental research fields. The use of pSEUDO::nisRK or pSEUDO-IL::nisRK circumvents the use of combinations of vectors in strains that do not carry nisRK on the chromosome. In those cases, pNZ9530 (18) is usually required to provide nisRK in trans, for complementation analysis. *L. lactis* JP9000 is preferable over the standard strain NZ9000, since pepN is intact in the former whereas it is disrupted in NZ9000 (pepN::nisRK) (18) (Figure 2). *L. lactis* IL9000 is the first IL1403 derivative in which the NICE system for nisin-inducible gene expression can be employed. The general aminopeptidase N (PepN) (21), while generally assumed not to be relevant under most conditions, is likely to have an effect on nitrogen metabolism through the influence of the products of peptide hydrolysis on e.g., the major pleiotropic regulator CodY (22). Also, the availability of specific peptides in the medium has been correlated with the ability of *L. lactis* to overproduce membrane proteins (23).
A) Expression of the membrane protein BcaP-GFP-His6 was induced in L. lactis strains NZ9000 and JP9000, both carrying plasmid pNZbcaP-GFP-H6 (11). The strains were grown in GM17 until an OD at 600 nm of 0.5 was reached, after which they were induced for one hour with 5 ng/ml of nisin. Mean fluorescence, as measured by flow-cytometry, of the plasmid-carrying strains is plotted, normalized to that of plasmid-free L. lactis NZ9000. The uninduced bar depicts the fluorescence of the non-induced JP9000 (pNZbcaP-GFP-H6) culture.

B) PepN activity in L. lactis strains MG1363 (24), NZ9000 (18), MG1363ΔpepN (25), and JP9000 (this work) was determined in cell-free extracts of cultures grown in GM17 until an OD at 600 nm of 0.5, by monitoring the hydrolysis of the PepN substrate lysyl-p-nitroanilide, as described previously (26). (A) and (B): data are the average of 4 biological replicates and the error bars are the associated standard deviation.

Figure 2 - Heterologous protein production and activity of aminopeptidase PepN in L. lactis. (A) Expression of the membrane protein BcaP-GFP-His6 was induced in L. lactis strains NZ9000 and JP9000, both carrying plasmid pNZbcaP-GFP-H6 (11). The strains were grown in GM17 until an OD at 600 nm of 0.5 was reached, after which they were induced for one hour with 5 ng/ml of nisin. Mean fluorescence, as measured by flow-cytometry, of the plasmid-carrying strains is plotted, normalized to that of plasmid-free L. lactis NZ9000. The uninduced bar depicts the fluorescence of the non-induced JP9000 (pNZbcaP-GFP-H6) culture. (B) PepN activity in L. lactis strains MG1363 (24), NZ9000 (18), MG1363ΔpepN (25), and JP9000 (this work) was determined in cell-free extracts of cultures grown in GM17 until an OD at 600 nm of 0.5, by monitoring the hydrolysis of the PepN substrate lysyl-p-nitroanilide, as described previously (26). (A) and (B): data are the average of 4 biological replicates and the error bars are the associated standard deviation.
To facilitate the integration of promoter-gfp reporters in the chromosome of L. lactis, pSEUDO-GFP was also constructed (Figure 4). This vector was made by cloning the gene of the superfolder variant of GFP (27) together with the two terminator sequences from the iGEM Biobrick I746909 (http://partsregistry.org/Part:BBa_I746909) (28) to block read-through into llmg_0576, a putative transcriptional regulator of the TetR family. The PCR fragment obtained with primers GFP-SF_Forw and GFP-SF_Rev was inserted in pSEUDO using BamHI and XhoI. The first primer was used to add 4 well-translated codons and a strong RBS sequence to the 5'-end of gfp (29).

As an example of the applicability of pSEUDO-GFP, promoter-gfp fusions were constructed using the upstream regions of two general stress response genes, hrcA and groES. The promoter region of hrcA was amplified using the primer pair PhrcA_forw/PhrcA_rev and that of groES was amplified with the primers PgroES_forw/PgroES_rev. EcoRI and XhoI were used to digest these DNA fragments and pSEUDO-GFP, in which the two promoter fragments were separately inserted. Insertion of the promoter-gfp fusions in the llmg_pseudo_10 locus of the chromosome of L. lactis NZ9000 yielded the strains NZ9000::PhrcA-gfp and NZ9000::PgroES-gfp. By exposing both strains to high temperatures, and making use of the stress-induced activity of P_hrcA and P_groES, it was demonstrated that they are able to reliably monitor, in a non-invasive manner (unlike e.g., promoter-lacZ fusions) and in real-time, the effect of growth and environmental conditions on the general fitness of cells (Figure 3).

The use of pSEUDO-GFP enables applying high-throughput screening strategies (e.g., using microtiter plates) for conditions in which putative promoters are expected to be active. Furthermore, the use of flow-cytometry, or other single-cell analysis techniques, allows characterizing phenomena such as gene expression heterogeneity.

Altogether, pSEUDO, pSEUDO-GFP and the derived plasmids and strains add value to the lactococcal research community in that they establish an improved working standard for the effective and efficient integration of DNA fragments into the chromosome of L. lactis.
Figure 3 – Induction of GFP production in *L. lactis* NZ9000::P\_groES-GFP and in *L. lactis* NZ9000::P\_hrcA-GFP, grown in GM17 medium at 30°C. Both cultures were shifted to 60°C when they had reached an optical density at 600 nm of 0.5. White: 0 min after the temperature shift; light grey: 15 min after the temperature shift; dark grey: 60 min after the temperature shift. The fluorescence was measured over time using an Epics XL-MCL flow-cytometer (Coulter, Fullerton, CA). Values were normalized to the fluorescence of *L. lactis* NZ9000 undergoing the same heat treatment. 20,000 cells were measured per experiment and 4 biological replicates obtained per strain and per time point. The error bars are the associated standard deviations.

**Part II: Genomic Array Footprinting in Lactococcus lactis**

*Construction of pGh8T7:ISS1*

The ISS1 element from pGh8:ISS1 (30) was removed by digestion of the plasmid with *Eco*RI and *Hind*III and replaced by a *Hind*III-, *Eco*RI-digested PCR product of the ISS1 element generated from pGh8:ISS1 with primers ISS1\_T7\_up and ISS1\_T7\_down (12). This resulted in pGh8T7:ISS1, which, like pGh9T7:ISS1 (12), contains the ISS1 element with the T7 promoters on either of its sides (Figure 4).
Table 1 – Oligonucleotides used in this study. Restriction enzyme sites are underlined.

<table>
<thead>
<tr>
<th>Name</th>
<th>Nucleotide Sequence; 5' &gt; 3' $</th>
<th>Restriction enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1_pseudo10</td>
<td>GCCTAGACAATTGTCCCATAGCTTTGATCC</td>
<td>BglII</td>
</tr>
<tr>
<td>P2_pseudo10</td>
<td>GCGGATCCCTGCTTTTGAATTTGAAGGTTTTGAAAG</td>
<td>BamHI</td>
</tr>
<tr>
<td>P3_pseudo10</td>
<td>CGAATTCGGGCTGCTTGGATTAATGATTTG</td>
<td>EcoRI</td>
</tr>
<tr>
<td>P4_pseudo10</td>
<td>GCAGTACCAATTGAAGACAAAGAAAACC</td>
<td>KpnI</td>
</tr>
<tr>
<td>P1_yfjF</td>
<td>GCAGAAATCTAGACTTCAACATACAGAGACCTCG</td>
<td>XbaI</td>
</tr>
<tr>
<td>P2_yfjF</td>
<td>GCAGAAAATCTAGCTTTAGCTTTAGGGTTGAAAG</td>
<td>BamHI</td>
</tr>
<tr>
<td>P3_yfjF</td>
<td>GCAGAAATTCTAGCTTTAGCTTTAGGGTTGAAAG</td>
<td>BamHI</td>
</tr>
<tr>
<td>P4_yfjF</td>
<td>GCAGAAGCTAGACAAATCGAAGAAAGAAAGCC</td>
<td>KpnI</td>
</tr>
<tr>
<td>nisRK_Forw</td>
<td>TAAAGGGATCCGCTTTAGATCAGATAAAGGTACGG</td>
<td>BamHI</td>
</tr>
<tr>
<td>nisRK_Rev</td>
<td>AGATTGGATCCAAACTGATATCTTTGAGCAGCTTC</td>
<td>BamHI</td>
</tr>
<tr>
<td>GFP-SF_Forw</td>
<td>ATAGTCTCGAGTAAGGGCAATATGGAACATCTTTGCTTAAGGCAAGACTGC</td>
<td>XhoI</td>
</tr>
<tr>
<td>GFP-SF_Rev</td>
<td>ACTATGATCGCTTAAAGGCGAAGCCACCCATCGCG</td>
<td>BamHI</td>
</tr>
<tr>
<td>PhrcaA_forw</td>
<td>ATCTGGGATGGATCAGAATGATCTATTTATACAG</td>
<td>EcoRI</td>
</tr>
<tr>
<td>PhrcaA_rev</td>
<td>GCCTCGGATCCCAGATCGCTTTATTTATGACACTCT</td>
<td>XhoI</td>
</tr>
<tr>
<td>PgroES_forw</td>
<td>GGAACGAAATTCAGATGCTGGAGCTCCCTTTCTG</td>
<td>EcoRI</td>
</tr>
<tr>
<td>PgroES_rev</td>
<td>ATACTCTCGAGATTTTATATTTATACAG</td>
<td>XhoI</td>
</tr>
</tbody>
</table>

Figure 4 – Schematic representation of pGh8T7:ISS1 (this work) and pGh9T7:ISS1 (12). Both plasmids carry a repA (ts) gene encoding a temperature sensitive replication protein that is non-functional above 37°C (30, 31). The ISS1 insertion sequence is flanked by two outward-facing T7 promoters (12). pGh8T7:ISS1 contains the tetracycline resistance marker from pT181 (30, 32) while pGh9T7:ISS1 carries the erythromycin resistance marker from pIL253 (15, 30).
The pGh8T7:ISS1 replicon, like that of pGh9T7:ISS1 (12), is a thermosensitive derivative of pWV01 (33), in which repA contains four mutations compared to that of pVW01 (31, 34). In L. lactis these plasmids replicate normally up to 28°C but not above 37°C (30, 31). The Ts plasmids can be used to perform insertional mutagenesis not only in L. lactis but also in other lactococci, enterococci and streptococci, where ISS1 has been shown to transpose randomly throughout the genome (30).

The characterization of mutants that are pulled out during a given screening strategy can be performed by amplifying the chromosomal regions that are adjacent to where the plasmid has integrated. The transposed structure is flanked by ISS1 on each side (30), so the outward-facing T7 promoters can be used to produce local RNA fragments which pinpoint the site of integration and therefore allow identifying the genes that have been disrupted or transcriptionally affected.

**Genomic Array Footprinting in Lactococcus lactis**

In GAF, an entire transposon library can be screened, for example, for mutants that disappeared or became differentially more prevalent under a certain experimental setup by comparing the obtained mutant bank (target) to the control or original library (12). This is accomplished by digesting chromosomal DNA from each of the two populations (target and control), and amplifying RNA from the DNA segments that remain adjacent to the transposon insertion sites using the T7 RNA polymerase. The RNA is then used as a template to yield cDNA that is labelled and hybridized onto a DNA microarray (12).

To validate this strategy in L. lactis, we used the method as described before (12), with small libraries (≈ 100 CFU’s). Following the transformation of L. lactis with pGh9T7:ISS1 and the subsequent temperature up-shift to select for integrants, 3 small libraries of 32 CFU’s each were made. From these 3 libraries, one was obtained from a selection of small colonies, the second from medium sized ones, and the third from colonies larger than the average wildtype L. lactis colony. Additionally, 10 random colonies (3 small, 4 medium, 3 large ones) were independently picked. We selected and independently expanded in liquid culture, not only each of these 10 colonies, but also each library of 32 CFU’s, to dismiss in this validation stage any artifacts, such as overrepresentation bias, that might derive from growth rate differences among mutant strains.

Chromosomal DNA from each of the small transposon libraries and from each of the 10 randomly selected strains was independently purified. Then, two chromosomal DNA mixtures were prepared: DNA pool 1 contained equal amounts of DNA from each of the three small libraries (thus representing 96 CFU’s); the second preparation (DNA pool 2) contained this same DNA mixture to which chromosomal DNA from each of the 10 independently picked colonies was proportionally added (in total about ≈ 30% compared to the DNA from
each of the three small libraries). This setup served to mimic a situation where, during a selective process, 10 mutants had disappeared from the transposon library pool. We only used small libraries, of 96 and 96+10 CFU’s, to minimize the probability of a given insertion locus being represented in both the “96 CFU’s” and the “10 CFU’s” partial libraries of DNA pool 2.

DNA from each of these two mixtures was digested with AluI, which cuts the *L. lactis* chromosome on average every 233 bp. The “MEGAscript® T7 Transcription Kit” (Ambion, Texas, USA) was used to amplify regions adjacent to the T7 promoters, according to manufacturer instructions. The obtained RNA was purified with the “High pure RNA isolation kit” (Roche Molecular Biochemicals, Mannheim, Germany). Synthesis of cDNA and DNA microarraying were performed as described before using SuperAmine glass slides (ArrayIt, Sunnyvale, CA) spotted with duplicates of around 2500 ORF amplicons of *L. lactis* subsp. cremoris MG1363 (17, 35).

Analysis of the DNA microarray data indicates that, indeed, some spots on the DNA microarray slide display greater signal intensities, corresponding to the 10 extra mutants in the 96+10 CFU’s DNA sample, as compared to the 96 CFU’s library (Figure 5).

![Figure 5 – DNA microarray output of GAF using an *L. lactis* transposon library of 96 CFU’s (horizontal axis) versus 96+10 CFU’s (vertical axis). Spots with significantly greater intensities in the channel that corresponds to the 96+10 CFU’s sample are highlighted as red triangles.](image-url)
As previously observed (30), integration of pGh9T7:ISS1 in the genome of *L. lactis* appears to have been random since the DNA microarray spots with greater intensities and those with greater intensity biases (Figure 5) map onto regions all around the genome (Table 2). Also, Southern hybridization on the DNA of all 10 independently picked CFU’s using pGh9T7:ISS1 as a probe indicates that integration *loci* were different in all cases as the samples yielded different patterns on the blot (data not shown).

Note, however, that the outcome from GAF depends on the restriction enzyme used to digest the chromosomal DNA. For the transposon insertion site to be identifiable, a DNA microarray probe has to at least partially match the chromosomal region between the T7 promoter and the restriction site of the particular restriction enzyme.

### Table 2 – DNA microarray probes that are highlighted as red triangles in Figure 5.

<table>
<thead>
<tr>
<th>Gene Microarray Probe</th>
<th>Net Intensity</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>96 CFU’s</td>
<td>96+10 CFU’s</td>
</tr>
<tr>
<td>lmg0430_cstA_MG752</td>
<td>431</td>
<td>773</td>
</tr>
<tr>
<td>lmg0724_lmg0724_mg10050s</td>
<td>336</td>
<td>825</td>
</tr>
<tr>
<td>lmg1031_trpC_MG750</td>
<td>352</td>
<td>8.543</td>
</tr>
<tr>
<td>lmg1032_trpD_MG749</td>
<td>322</td>
<td>1.653</td>
</tr>
<tr>
<td>lmg1453_lmg1453_MG150179</td>
<td>319</td>
<td>1.089</td>
</tr>
<tr>
<td>lmg1454_lmg1454_mg6115</td>
<td>253</td>
<td>8.020</td>
</tr>
<tr>
<td>lmg1455_bglA2_mg6114</td>
<td>275</td>
<td>3.149</td>
</tr>
<tr>
<td>lmg1455_bglA2_MG809</td>
<td>150</td>
<td>510</td>
</tr>
<tr>
<td>lmg1668_lmg1668_MG150107</td>
<td>349</td>
<td>2.830</td>
</tr>
<tr>
<td>probe mg5799 (between lmg1668 and lmg1669)</td>
<td>169</td>
<td>16.028</td>
</tr>
<tr>
<td>lmg1669_lmg1669_mg5800</td>
<td>173</td>
<td>6.440</td>
</tr>
<tr>
<td>lmg2100_ps442_MG1927</td>
<td>280</td>
<td>8.503</td>
</tr>
<tr>
<td>lmg2288_ackA2_MG1320</td>
<td>296</td>
<td>1.902</td>
</tr>
</tbody>
</table>

**Further Considerations on GAF**

The first step in the implementation of a GAF strategy is the construction a transposon library in the organism of interest. The choice of either pGh8T7:ISS1 or pGh9T7:ISS1 is important if the experimental setup requires the use of either erythromycin or tetracycline to select for a marker, other than that on any of those two plasmids.

The size of the library, that is the number of integrants to be screened, is also a matter of optimization. Assuming that plasmid integration via ISS1 is entirely random, the probability of a gene having been disrupted at least once in a given library is given by the formula:
\[ P = 1 - \left(1 - \frac{1}{N}\right)^n \]

where \(N\) is the number of genes on the chromosome and \(n\) is the number of CFU’s that were collected to make the library. Thus, a transposon library of 5.000 CFU’s of \(L.\ lactis\) MG1363, which encodes 2.563 genes (8), is likely to have a coverage of 85.8% disrupted genes represented at least once in that library. Please note however that the formula offers a simple estimate since 1) some genes are essential and will never be represented in any library, 2) gene sizes differ considerably and, for example, 3) a gene in an operon can be indirectly disrupted if transcription is terminated before reaching it.

A library made from a larger number of CFU’s is not necessarily better. Slight variations in the site of integration, for example of 100 bp, can result in different phenotypes but produce the same outcome on a DNA microarray slide when GAF is employed. Conversely, the opposite is also true. Thus, the disappearance of a given mutant under selective pressure could be masked by the persistence of other mutants, that are phenotypically distinct, but that yield similar signals using GAF. To minimize this effect, the selection of a proper restriction enzyme to digest the chromosomal DNAs of target and control cultures is essential. A non-frequently cutting enzyme produces on average large fragments, exacerbating the effect, *i.e.* making it more difficult to distinguish insertion sites using GAF. Although a more frequently cutting enzyme would enable greater resolution it will also increase the chance that there will be no probe on the DNA microarray that matches the region between the T7 promoter and the restriction enzyme site.

We therefore suggest including variations in the size of the transposon libraries and in the restriction enzymes used, which would also provide the replicates that are required in any experimental setup. For \(L.\ lactis\) one could consider using the restriction enzymes Alul, TaqI and Ddel, which on average cut the chromosomal DNA every 233 bp, 531 bp and 448 bp, respectively. Also, one should keep in mind that in GAF, the probe on the DNA microarray slide does not match one-to-one to the gene that has been disrupted, but rather to a region that lies between the transposon insertion site and the enzyme restriction site.

All that is discussed above is not only true for other lactococci, but also for enterococci and streptococci, where the GAF methodology presented here can be equally implemented since IS51 has been shown to transpose randomly throughout the genomes of these organisms (30).
Upgrading the Toolkit of Lactococcus lactis

References


