# Supporting Information:

# Effect of genome position on heterologous gene expression in *Bacillus subtilis*; an unbiased analysis

Christopher Sauer[1, 2], Simon Syvertsson[1], Laura C. Bohorquez[3], Rita Cruz[1, 2], Colin R. Harwood[1], Tjeerd van Rij[2], Leendert W. Hamoen[1, 3]

[1] Centre for Bacterial Cell Biology, Newcastle University, Newcastle upon Tyne, United Kingdom

[2] DSM Biotechnology Center, P.O. Box 1, 2600 MA Delft, The Netherlands

[3] Bacterial Cell Biology, Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, the Netherlands

## Contents

**Table S1**

Transposon insertion efficiency of the expression cassette from plasmid pSS125. The mean efficiency was low, a 1.2:1 ratio of kanamycin to erythromycin resistance was observed, compared with that of the pMarB containing the original mariner transposon (29.1:1). Maximal, minimal and mean values are given for the insertion efficiencies of several (n) experiments.

| | **$Kan^R$ : $Erm^R$ ratio** | | |
|---|---|---|---|
| **Plasmid** | **Maximum** | **Minimum** | **Mean** |
| pMarB (n = 6) | 45:1 | 14:1 | 29.1:1 |
| pSS125 (n = 4) | 1.4:1 | 0.9:1 | 1.2:1 |

**Table S2**

Primers used in this study.

| Name | Sequence |
|---|---|
| oSS258 | gtggatgcaatgggtaccctgcagatgag |
| oSS259 | cagggtacccattgcatccacctcactac |
| oSS260 | caattccacacattatgccacaccttgtagataaagtcaacaaccgtaatacgctgacaagagag |
| oSS261 | gtggcataatgtgtggaattgtgagccgctcacaattatatagggaaaaggtggtgaac |
| oSS262 | caatgagctgggttttttgtttgttaattattatttttgacaccagac |
| oSS330 | ggttcaattggcctacgaggaatttgtatc |
| oSS331 | ggttcaattgggacccctatctagcgaac |
| oSS344 | gcggagctctaactgcgccaggtgcagttg |
| oSS345 | gcggagctctaagtgcttccaggaatcgtcg |
| oSS346 | gcaggaattcggatccttattatttgtagagctcatccatgccatg |
| oSS347 | gcaggaattcggatccattattatttttgacaccagacc |
| oSS348 | ggatcgcatgctaattaacagcggggctgttcg |
| oSS349 | ggatcgcatgctaacctatgatttttcgttagatcc |
| Arb1 | ggccacgcgtcgactagtacnnnnnnnnnngatat |
| Arb2 | ggccacgcgtcgactagtac |
| MarB1 | gttagaccggggacttatc |
| MarB1N | gcgcctacgaggaatttgtatc |
| MarB2 | gcatttaatactagcgacgcc |
| MarB2N | gggaatcatttgaaggttgg |
| OriC fwd | gaattccttcaggccattga |
| OriC rev | gatttctggcgaattggaag |
| Ter fwd | tccatatcctcgctcctacg |
| Ter rev | attctgctgatgtgcaatgg |
| Loc2 fwd | caggacatgttccaggtatc |
| Loc2 rev | gacctcaagcgtaatgactg |
| Loc3 fwd | gaagccgattccaaggatac |
| Loc3 rev | agcttatcagagccagagtc |
| Loc4 fwd | gctatgtcaggcgataaacc |
| Loc4 rev | ttcaacaggtctgcctgttc |
| Loc5 fwd | gcagacagcctgttcaattc |
| Loc5 rev | gcgtgcagatcgtctaatgg |
| Loc6 fwd | caggcctcattgccttaaac |
| Loc6 rev | cttagcagcggcaacgtttc |
| Loc7 fwd | tcaaaggcaacgggccttac |
| Loc7 rev | aatagccagtttaccgcccg |
| Loc8 fwd | agaatacatgcgccagcttg |
| Loc8 rev | tacaagagttgtgccctctg |
| Loc9 fwd | gaacggaccgaatgttaacc |
| Loc9 rev | actggaagaaagtgagctgg |
| Loc10 fwd | aacacatgaagcggtcgaac |
| Loc10 rev | ggtcgagctggtagatttcc |
| Loc12 fwd | gagactgctgcaccttaatg |
| Loc12 rev | ccgatgcagtcaatgacgtg |
| Loc13 fwd | aagagtttgccgcgtttgtc |
| Loc13 rev | gtaagcggagaggtaatgac |
| Loc14 fwd | aagtggcgcccgatcattac |
| Loc14 rev | cgcaagcaatcagccaagtg |

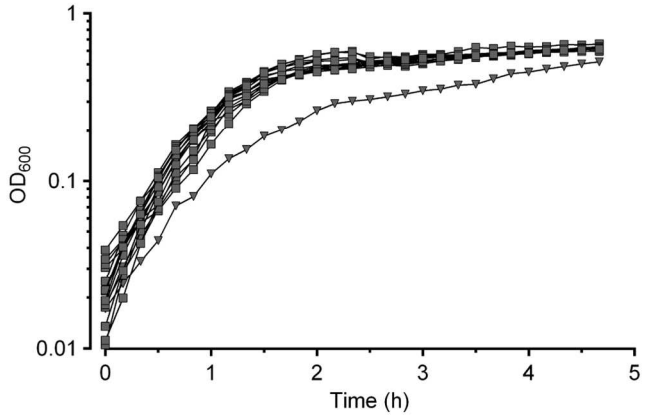


**Figure S1.** Growth rate measurement of transposons. Square symbols represent strains used in further experiments. One transposon strain showed a clear growth rate defect (triangles) and was discarded. Average values from three technical replicates are shown.

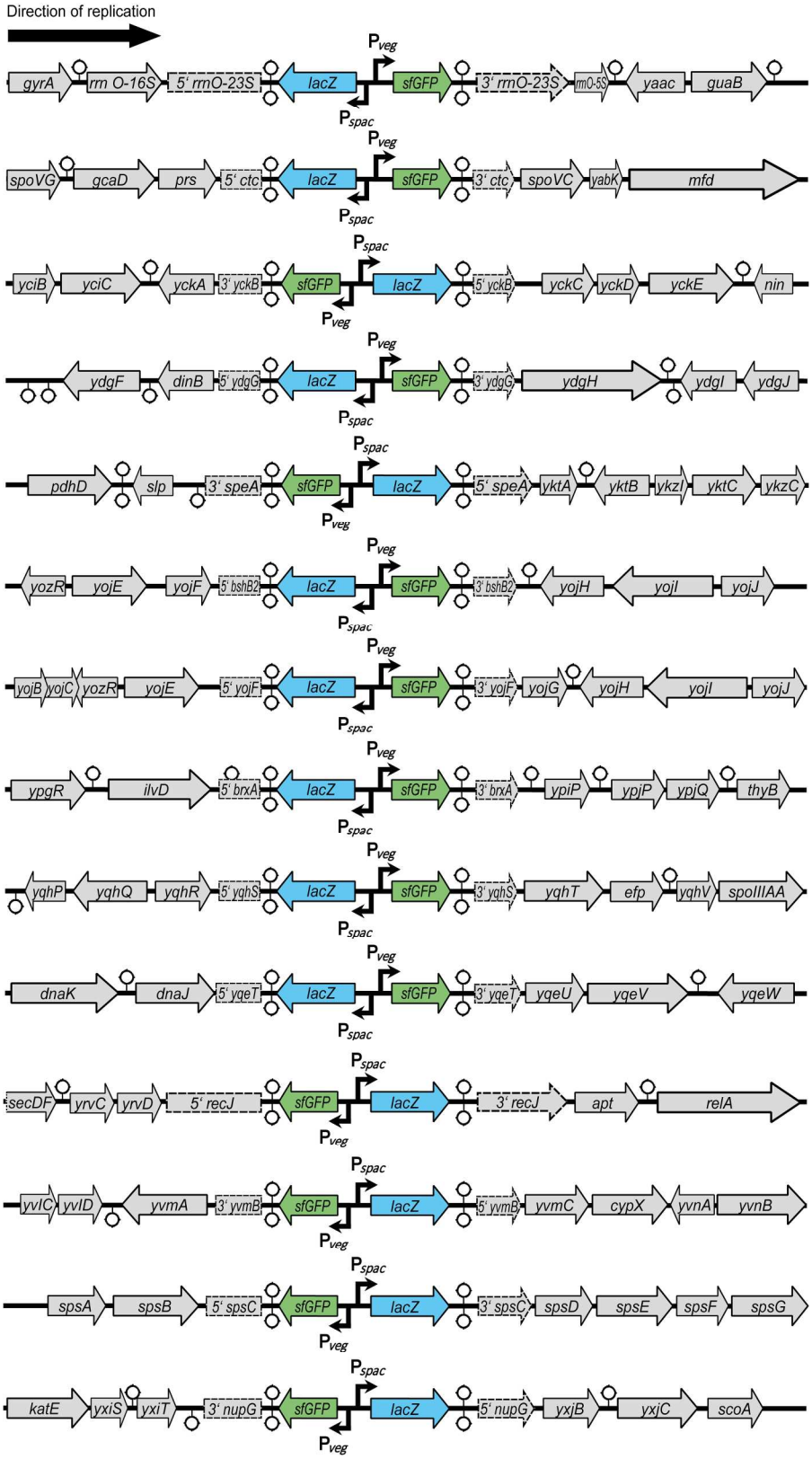


**Figure S2.** Overview of the chromosomal transposon insertion sites. Dotted lines show genes disrupted by the transposed expression cassette. The large black arrow indicates direction of DNA replication.

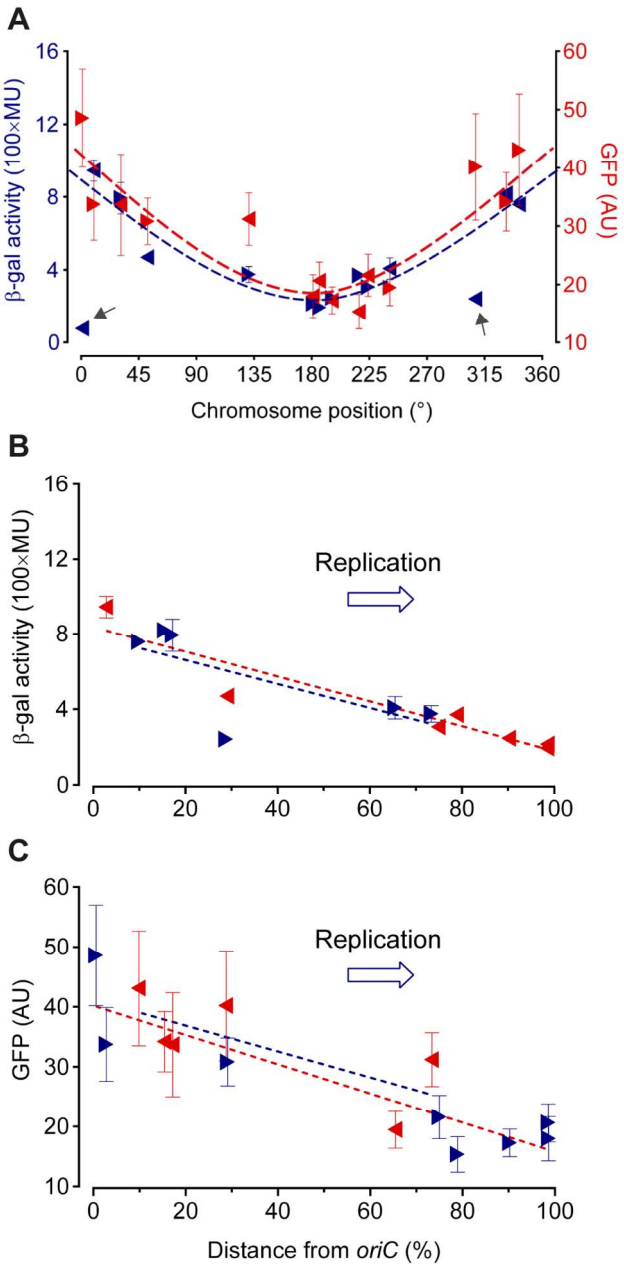


**Figure S3.** Effect of chromosomal location and transcription direction on gene expression. Cells were grown in LB medium at 37°C and samples were taken during exponential growth ($OD_{600}$ ~ 0.5). (A) β-galactosidase activities (blue) and GFP fluorescence (red) are plotted against chromosomal location (degrees) of the transposons. Outliers are marked by grey arrows. Goodness of fit for LacZ with $R^2$ = 0.79 (excluding locus at 1°) and GFP with $R^2$ = 0.79. β-galactosidase activities (B) and GFP fluorescence (C) are plotted against the distance of the transposon insertion sites from the origin of replication (*oriC*). Reporter genes that are transcribed in the same direction as DNA replication are shown in blue and genes that are transcribed against the DNA replication direction are shown in red. Triangles indicate transcriptional direction of reporter genes. Error bars indicate the standard deviation

of two technical replicates in case of β-galactosidase and at least 100 cells for GFP. Goodness of fit for LacZ with $R^2 = 0.48$ (blue), $R^2 = 0.87$ (red) and for GFP with $R^2 = 0.81$ (blue), $R^2 = 0.52$ (red).

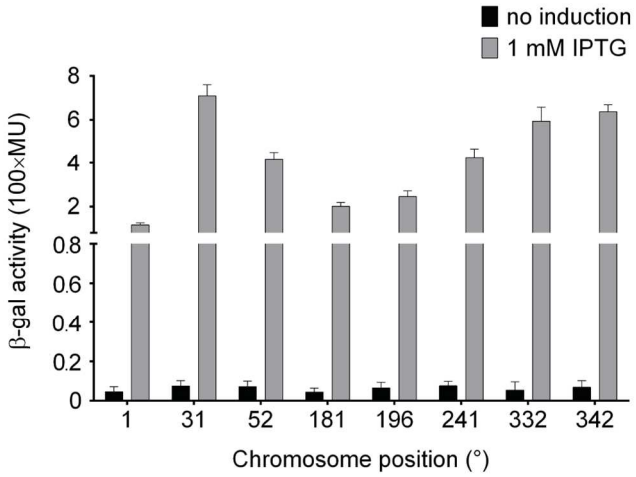


**Figure S4.** Assessing transcriptional read-through into *lacZ* by measuring β-galactosidase activities in the absence of the inducer IPTG in stationary phase after 3.5 h of growth. Eight transposon strains spread out over the genome were assayed under non-inducing conditions (black bars) and when grown in the presence of 1 mM IPTG (grey bars). Average values with standard deviations of three biological replicates are shown.

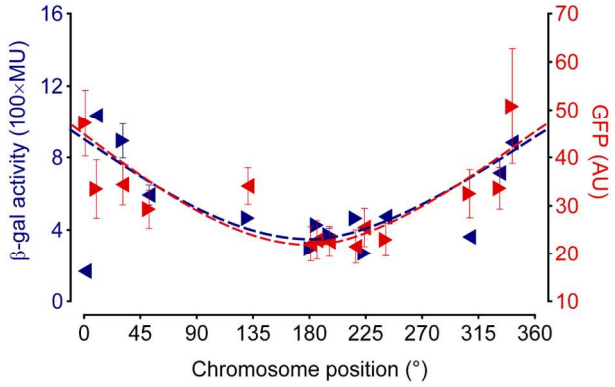


**Figure S5.** Expression differences in stationary growth after 3.5 h of growth. β-galactosidase activities (blue) and GFP fluorescence (red) are plotted against chromosomal location (degrees) of the transposons. The direction of the triangles indicates the transcriptional direction of *lacZ* or *gfp*. Representative results are shown from a single experiment. Error bars indicate the standard deviation of two technical replicates in case of β-galactosidase and at least 100 cells for GFP. Goodness of fit for LacZ with $R^2 = 0.82$ (excluding locus at 1°) and GFP with $R^2 = 0.70$.

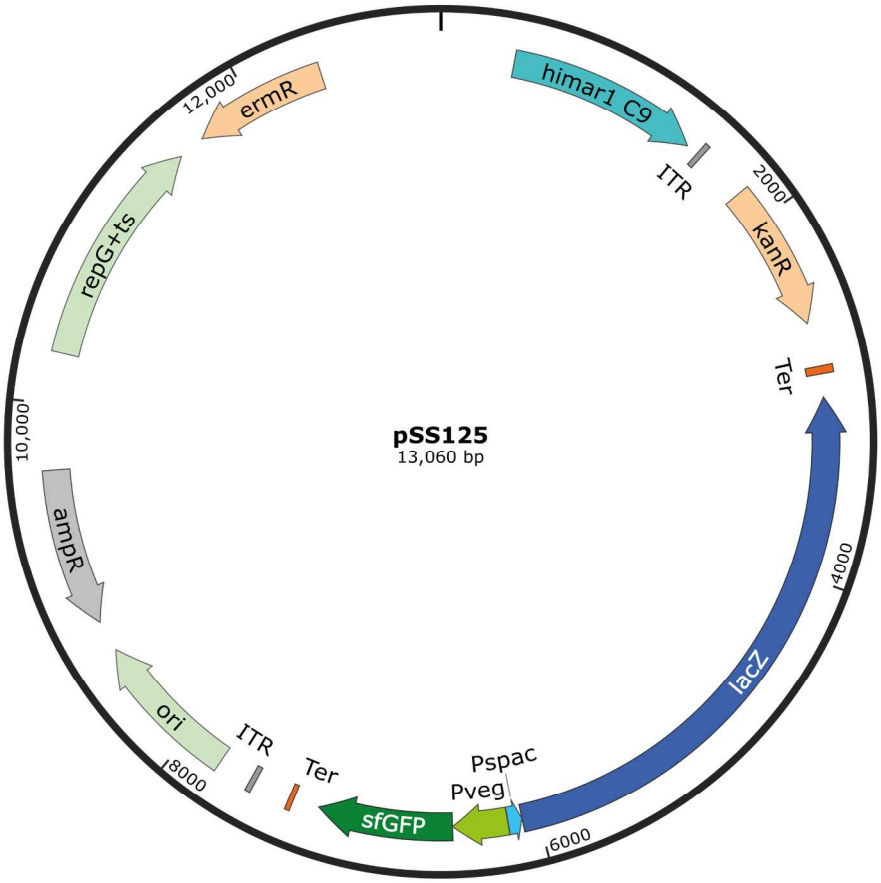


**Figure S6.** Plasmid map of the shuttle vector pSS125 containing the *himar1* transposase gene (cyan), ITRs (grey), kanamycin resistance gene (light orange), transcriptional terminators (orange), $P_{spac}$ driven *lacZ* (blue) and $P_{veg}$ driven *sfGFP* (green).