Final report

Improving bacterial metabolic efficiency and genomic stability by directed genome shuffling.

OTKA Identifying number 106231

Principal investigator: Kinga Umenhoffer

Summary

Genome-scale engineering now reached a state, where it enables massive reorganization of microbial genomes, including assembly of semi-synthetic bacterial genomes[1]. Using synthetic biology approaches, our group has pioneered experimental genome reduction of *E. coli* by systematically deleting laterally transferred (LT) islands (prophages, mobile elements) from the genome of *E. coli* K-12[2]. Taking the advantage of this pioneering work, and the robustness of Multiplex Automated Genome Engineering[3] (MAGE) in combination with CRISPR-guided genome editing[4], we proposed to construct genomic hybrid E. coli strains. Our aim was to investigate the evolutionary role of prophages[5,6] and other mobile genetic elements in the course of evolution, to discover the limits of re-arranging a genome, and produce an improved, genomically stabilized, industrially important E. coli BL21 strain. As a result, we have constructed a hybrid bacterial genome by the removal of destabilizing prophage sequences via genome shuffling, and by subsequent mass-inactivation of all mobile genetic elements. The latter step represents a novel technical approach, by which the mobile insertion elements can be inactivated without a priori knowledge of their exact genomic location and directional information. The project is ready for publication, the manuscript is under preparation and the resulting strain is already under consideration for industrial applications by Scarab Genomics LLC, Madison, USA. Patenting of the strain is now in a negotiation phase with the company.

Design of deletions and initial modifications.

Our primary goal was to identify the possible deletion targets of BL21DE3 genome. We used the multi-genome aligner Mauve program to align the available genome sequences of BL21DE3, MG1655 and MDS42. Genomic comparisons were carefully analysed manually using the Seqbuilder package provided by DNASTAR Inc. Genomic comparisons resulted regions harboring 9 prophages and 11 different types of intact IS elements in 62 sites in BL21DE3. The original BL21 strain harbours only 8 prophages[7], but prophage (DE3) was inserted into the genome to maintain inducible expression of the highly selective T7 RNA polymerase system[8] (controlled by an IPTG inducible *lacUV5* promoter). 1. The genomic region containing the prophage DE3 has no sequence homology in our K-12 strain collection, and cannot be easily deleted by phage transduction, due to the lack of sequence homology around this genomic region. This particular segment was removed from the genome by suicide plasmid-mediated deletion, as a first step[9]. 2. From one of our special K-12 strains we transferred a special T7

RNA polymerase gene under the control of a modified lac promoter and operator into the modified, DE3-less BL21 cell line. 3. To further increase the very low transduction rate and to increase the stability of transforming DNA, we removed the genomic region responsible for host restriction- modification. Considering the strain specific island in this genomic segment, we had to construct specific, new marked region in K-12 for the P1 transfer. As a final result of this process, we successfully created a BL hybrid strain, in which the host restriction and modification system has been deleted, and the DE3 T7 polymerase-expressing system has been replaced by a chromosomally integrated, tightly controlled lac-T7 polymerase system.

Creating a set of single-deletion strains to remove prophages.

The genome of the original BL21 harbours 8 prophages from which six occupy similar genomic positions in both BL and K-12 strains[10]. We've aimed to remove them primarily to create a prophage-free hybrid strain. In order to remove these genome destabilising specific genetic elements from BL21, the corresponding chromosomal segments were marked with an antibiotic-resistance gene in the donor K-12 strains. Transductions of BL21 strain with P1 phages harbouring the desired fragments from the clean genomes were sequentially carried out to construct a set of single-deletion strains. Following the P1 transduction, clones from the recombinant strain pool were validated with site specific primers to identify the extent of genome exchange. Transductants with different genotypes from each engineering step were screened whether they retained efficient recombinant protein-expression and good growth profiles. The high-expression hybrid clones were used to create a prophage-free BL21 variant as a result of sequential P1 transduction followed by marker removal. Altogether, multiple, large regions of the BL genome were replaced by the corresponding K-12 MDS sequence, ending up with the elimination of all prophages from the targeted genome (Fig1). The resulting hybrid strain BL21D9 has a reduced genome (91%), and 6,7% (280 kb) of it originated from *E.coli* K-12.There are 280 genes affected, including a complete transfer of 6 unique K-12 genes and the elimination of host restriction-modification system (*ΔhsdR, ΔhsdM, ΔhsdS*) (*ΔmcrA, ΔmcrB, ΔmcrC, Δmrr*). Moreover, it carries an improved T7 polymerase-expressing module.

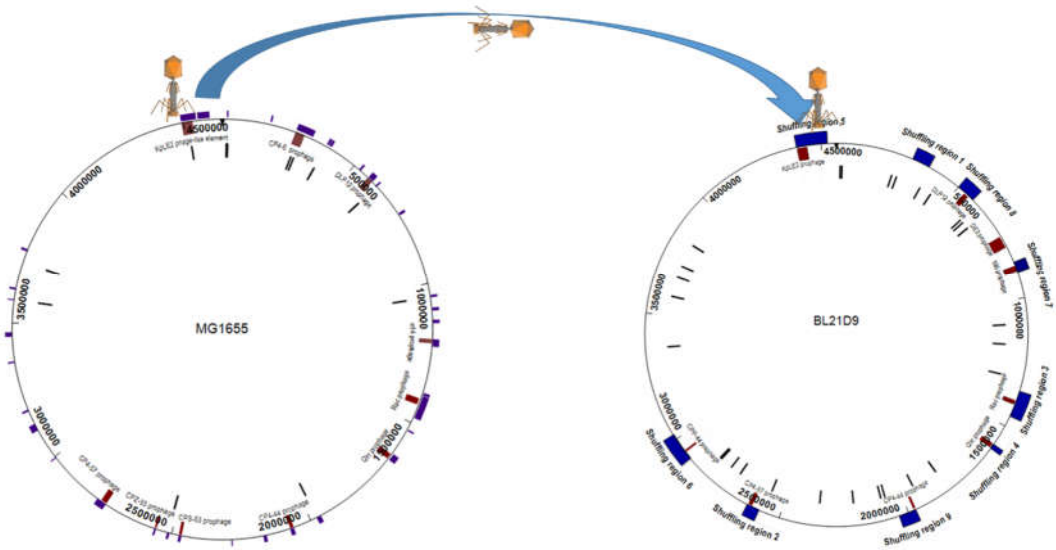

Figure 1: Creating a prophage-free BL strain by sequentially replacing prophage -harboring LT islands by the corresponding, marked segments of the "clean" MDS (K-12) genome.

Mass-inactivation of mobile genetic elements (insertion elements).

With the elimination of prophages, a decent amount of insertion elements was also removed from the hybrid genome. However, 30 copies of IS elements from 7 IS classes still remained, mosaicly distributed across the chromosome of BL21D9, on both replichores and in both directions. Instead of eliminating every single one of them with genome shuffling, we sought to find an efficient, less time-consuming and novel synthetic biological approach. Taking advantage of the robustness of the newly developed CRISPR/Cas9 aided genome editing and Multiplex Automated Genome Engineering (MAGE), we developed a novel workflow to carry out simultaneous mass-inactivation of endogenous transposases with a previously unprecedented speed. MAGE is an allelic replacement dependent genome editing technique, in which targeting ssDNA oligos are repeatedly introduced into a bacterial cell, simultaneously targeting and modifying multiple genomic locations. Structure and sequence of insertion sequence elements within each IS classes are highly conserved, this enables their mass targeting with MAGE, by introducing translational stop mutations (double stop codon + frame-shift = TAATAGG) into the corresponding transposase genes. All classes of IS elements on both strands with complemental oligonucleotides were targeted in subsequent MAGE cycles to overcome positional and directional information requirements (Figure 2).

IS type	copy number	MAGE oligo	Mage cycle	CRISPR/Cas
insertion sequence:IS1	17	sense/antisense	10/8	+
insertion sequence:IS150	4	sense/antisense	4/4	+
insertion sequence:IS186	4	sense/antisense	4/4	+
insertion sequence:IS3	2	sense/antisense	4/4	+
insertion sequence:IS30	1	sense	2/0	-
ISEcB1	1	sense	2/0	-
IS5	1	sense	2/0	-

Figure 2 Distribution and copy number of the remaining IS elements during mass inactivation

Next, we used CRISPR/Cas9 aided genome editing technique to enrich the multiple allelic-replacement events following MAGE cycles. By transforming the Cas9 protein-expressing plasmid into the MAGE-treated library populations and targeting it with the appropriate guide RNAs, we've successfully eliminated cells harbouring the wild type sequence of the transposase by site-specific dsDNA cleavage. Induction of double-strand genomic breaks rapidly kills cells in the presence of even a single wild-type transposase sequence, regardless of its genomic location (Figure 3). The resulting strain (BL21D9 ISKO) is a prophage-free, BL21 based hybrid strain that has all its mobile genetic elements inactivated. In this work Ákos Nyerges and Gábor Draskovits had contribution.

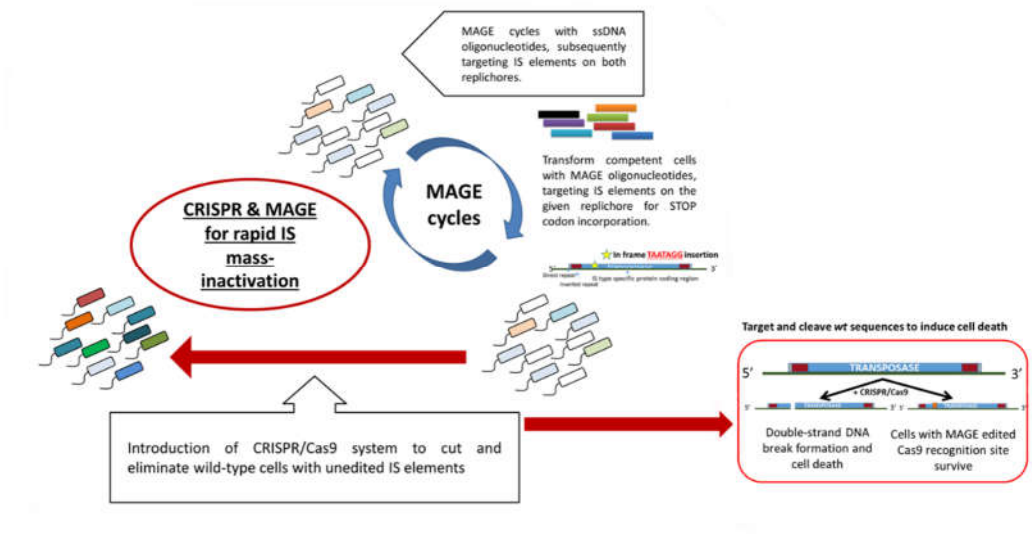

Figure 3. Mass-inactivation of mobile IS elements with a CRISPR/Cas9 aided genome editing and Multiplex Automated Genome Engineering combined method

Targeted inactivation of the major error-prone DNA polymerases.

In a previous work, removal of diversity-generating, error-prone DNA polymerase enzymes (*polB*, *dinB* and *umuDC* genes) involved in induced mutagenesis resulted in a significant stabilization of the genome[11]. To further stabilize the hybrid strains under stress conditions, we used the same inactivation technique on the genes of error-prone DNA polymerases. The resulting strain (BL21D9 ISKO_pdu) showed a significant attenuation in mutation rate under stress conditions. In this part of the project, Bálint Csörgő contributed to the project

Phenothypical analysis of the hybrid strains.

All prophage-free and IS inactivated hybrid strain lines were thoroughly analysed, regarding growth profiles, recombinant protein expression, metabolic profiles, and strain stability.

Measuring recombinant protein expression. Recombinant protein expression measurement using the widespread T7 induction system showed improved protein expression and tighter regulation in hybrid strains, compared to the wild type (Figure 4). In addition, comparative protein expression experiment, using an IPTG-inducible GFP expressing plasmid vector under the control of T5-lac promoter, show significant increase in the protein expression yield (based on final fluorescence yields).

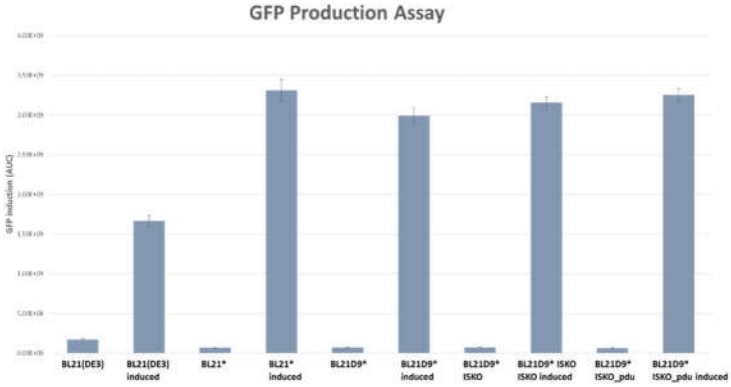

Figure 4 Recombinant protein expression was tested with a GFP-expressing plasmid driven by a T7 promoter.

Basic growth properties. Genomic modifications in all BL21 hybrid strains have minimal effect on growth properties in LB medium compared to wild type. More closely, the hybrid strain-set shows unaltered doubling time and lag phase, but reduced maximal OD when growing in minimal-salts (10mM glucose) medium. This property will be further investigated in the future, in order to reach improved fitness properties under industrially relevant growth conditions.

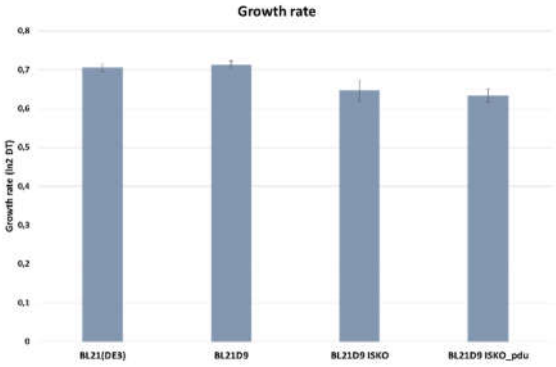

Figure 5 Growth properties of the hybrid BL strain series in LB medium

Metabolic profiling. In order to characterize metabolic capabilities of the prophage-free BL hybrid strain in relation to the wild-type strain, Phenotypic Microarray measurements were carried out using Plates (BioLog) PM1 and PM2A. Out of ~200 carbon sources, in the presence of only eight showed the hybrid strain significantly altered fitness properties ($P<0.05$, relative fitness<0.7 or >1.3 compared to BL21DE3). These metabolic alterations have minor practical effect under industrial conditions (Figure 6).

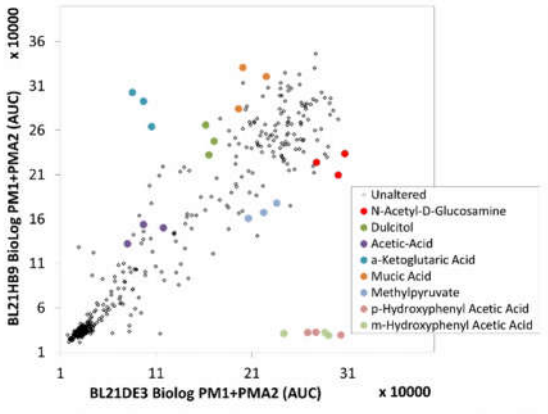

Figure 6 General metabolic activity of BL21HB9 and BL21DE3 was compared using phenotypic microarray assay plates (BIOLOG PM1 and PM2A).

Stress resistance. B strains are prone to lysis via the induction of its prophages (mainly DE3)[12]. Strain stability is substantially improved with the elimination of stress-induced prophage action, and the removal of λ-lysogen phenotype. The ease of engineering and propagating vectors slightly improved, due to the elimination of Endonuclease I (EndA-)[13] and to the eliminated host restriction and modification system. Epsilometer test (bioMerieux's E-test)

showed no changes in antibiotic susceptibility for various types of antibiotics, which also suggests improved vector stability.

Mutational assays. To detect the rate and spectrum of spontaneous mutations in BL hybrid lines, cells resistant to *d-cycloserine* were selected in fluctuation assays[14]. Wild type BL21(DE3) shows a mutation rate at least one order of magnitude higher than the K-12 strains. More importantly, over 75% of *cycA* mutations were IS insertions. Removal of the prophages from the genome, contrary to our expectation, did not affect the mutation rate, nor the distribution of mutation types. However, inactivation of the endogenous IS elements in the BL hybrid extinguishes the translocation of the element, and eliminated this portion of the mutation spectra. We found an 75% decrease in spontaneous mutation rate in BL21D9 ISKO. Mutation rates were also measured under stressful conditions with overproduction of a toxic protein (ORF238[15]). Targeted inactivation of the stress-induced error-prone DNA polymerases in the strain BL21D9 ISKOpdu facilitated further genome stability (Figure 7)

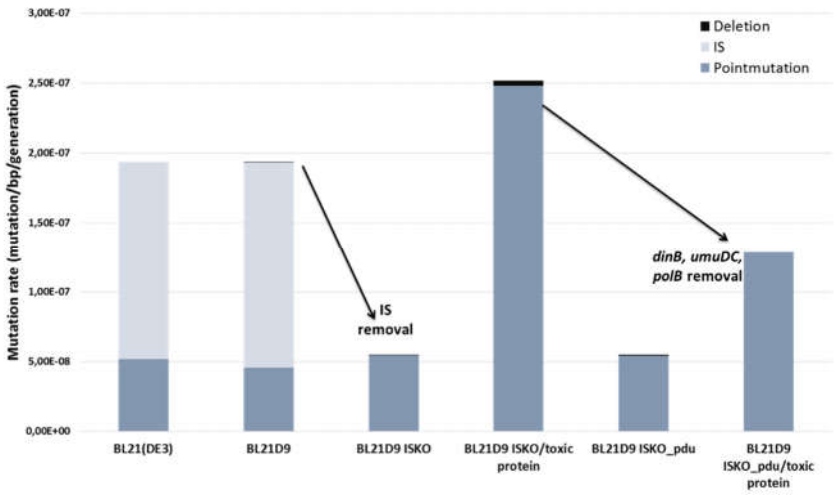

Figure 7. Mutation rate and spectra of mutations without stress and under toxic protein overexpression

Congress participations

FEBS EMBO 2014 (Paris) poster presented

Synthetic Biology Gordon Research Conference 2015 (Boston) poster

EMBO Meeting 2015 (Birmingham) poster, selected for flash talk presentation

EMBO | EMBL Symposium: The Mobile Genome: Genetic and Physiological Impacts of Transposable Elements 2015 (Heidelberg) poster presented

Manuscripts recently submitted (partially related to this project):

Ákos Nyerges, Balint Csorgo, István Nagy, Kinga Umenhoffer, Balazs Bogos, György Pósfai, Csaba Pal

pMAGE: a highly precise and portable genome engineering method

Ildikó Karcagi*, Gábor Draskovits*, Kinga Umenhoffer, Gergely Fekete, Orsolya Méhi, Károly Kovács, Zsuzsanna Györfy, Tamás Fehér, Balázs Bogos, Frederick R. Blattner, Csaba Pál, György Pósfai, Balázs Papp

An experimental test of the adaptive genome streamlining hypothesis

Manuscript based on this project is currently under preparation.

References

1. Pál, C., Papp, B. & Pósfai, G. The dawn of evolutionary genome engineering. *Nat. Rev. Genet.* **15,** 504–512 (2014).
2. Pósfai, G. *et al.* Emergent properties of reduced-genome Escherichia coli. *Science* **312,** 1044–1046 (2006).
3. Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460,** 894–898 (2009).
4. Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346,** 1258096–1258096 (2014).
5. Wang, X. *et al.* Cryptic prophages help bacteria cope with adverse environments. *Nat. Commun.* **1,** 147 (2010).
6. Bondy-Denomy, J. & Davidson, A. R. When a virus is not a parasite: the beneficial effects of prophages on bacterial fitness. *J. Microbiol.* **52,** 235–242 (2014).

7. Jeong, H. *et al.* Genome sequences of Escherichia coli B strains REL606 and BL21(DE3). *J. Mol. Biol.* **394,** 644–652 (2009).
8. Studier, F. W. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189,** 113–130 (1986).
9. Kolisnychenko, V. *et al.* Engineering a reduced Escherichia coli genome. *Genome Res.* **12,** 640–647 (2002).
10. Studier, F. W., Daegelen, P., Lenski, R. E., Maslov, S. & Kim, J. F. Understanding the differences between genome sequences of Escherichia coli B strains REL606 and BL21(DE3) and comparison of the E. coli B and K-12 genomes. *J. Mol. Biol.* **394,** 653–680 (2009).
11. Csorgo, B., Feher, T., Timar, E., Blattner, F. R. & Posfai, G. Low-mutation-rate, reduced-genome Escherichia coli: An improved host for faithful maintenance of engineered genetic constructs. *Microb. Cell Factories* **11,** 11 (2012).
12. Edlin, G., Lin, L. & Bitner, R. Reproductive fitness of P1, P2, and Mu lysogens of Escherichia coli. *J. Virol.* **21,** 560–564 (1977).
13. Dürwald, H. & Hoffmann-Berling, H. Endonuclease I-deficient and ribonuclease I-deficient Escherichia coli mutants. *J. Mol. Biol.* **34,** 331–346 (1968).
14. Fehér, T., Cseh, B., Umenhoffer, K., Karcagi, I. & Pósfai, G. Characterization of cycA mutants of Escherichia coli: An assay for measuring in vivo mutation rates. *Mutat. Res. Mol. Mech. Mutagen.* **595,** 184–190 (2006).
15. Umenhoffer, K. *et al.* Reduced evolvability of Escherichia coli MDS42, an IS-less cellular chassis for molecular and synthetic biology applications. *Microb. Cell Factories* **9,** 38 (2010).