Volume 59(Suppl.2):169-187, 2015
Acta Biologica Szegediensis
http://www.sci.u-szeged.hu/ABS

REVIEW

Microbial genome engineering for promoting health and understanding disease

Tamás Fehér*, Zsuzsa Györfy

Synthetic and Systems Biology Unit, Institute of Biochemistry, Biological Research Centre of the Hungarian Academy of Sciences, Szeged, Hungary

ABSTRACT The completion of the first microbial genomes nearly two decades ago opened a completely new chapter in molecular genetics. The availability of precise sequence data permitted the extended use of existing genome engineering methods, and urged the development of a novel set of more rapid and simple techniques for genome editing. The rapidly decreasing price of sequencing and DNA synthesis opened further possibilities of high-throughput genetic analysis and assembly. As a consequence, biomedical knowledge increased at an exponential rate and accelerated the development of numerous connecting fields, including that of medical microbiology. This review presents the reader the toolbox available today to edit and assemble microbial genomes and showcases the key molecular genetic strategies employed to dissect the mechanisms of pathogenesis and construct microbial strains for preventive or therapeutic applications.
Acta Biol Szeged 59(Suppl.2):169-187 (2015)

KEY WORDS

engineered probiotic bacteria
genome editing
genome synthesis
phage therapy
smart therapeutics

Introduction

The milestones of a scientific field are rarely identified at their time of occurrence. This is also the case for medical microbiology, a discipline which through its spectacular progress, has made one of the most important contributions to improving the expectancy and quality of human life. The prevention, diagnostics and therapy of microbial infections available today would be unimaginable without the knowledge cumulated in this field. Leuwenhoek's first description of microbial cells (bacteria, fungi and protozoa) in a raindrop using a microscope opened the way to studying microbial morphology, but the relation of bacteria to human diseases was to come nearly two centuries later. Pasteur's experimental evidence for biogenesis, and disapproval of spontaneous generation was an important support for the germ theories of fermentation and disease. Koch's postulates founded the golden age of discovering the causative agents of numerous infectious diseases. The discovery of penicillin by Fleming opened a whole new era of antibacterial treatment. In the second half of the twentieth century, the application of molecular genetic analysis in bacteria and bacteriophages granted a continuous progress within medical microbiology, paving the way for the exponential explosion of information and possibilities brought about by the first completed microbial genome sequences in 1997.

Despite being a recent event, the onset of whole-genome sequencing is generally considered to be a milestone within life sciences. The consequential acceleration in the growth of genomic sequence data introduced a revolution in at least three scientific subdivisions: i) evolutionary biology, where sequence data warranted the construction of phylogenetic trees of previously unseen detail; ii) systems biology, where the building of genome-scale models finally became a reality; and iii) molecular genetics, where the exact and complete DNA sequences permitted the targeted modification of any gene of choice to unravel its function. This review focuses on a subset of the third category, illuminating the results and possibilities of engineering microbial genomes with the aim of promoting human health. Prominent examples will be given of projects where bacteria or phages harboring precisely edited genomes are used to prevent or treat diseases, or at least have been shown to work in proof of principle experiments. To give the reader a complete picture, the most commonly used targeted genome editing and genome assembly techniques will be introduced first, with special attention to their application in a multiplex or high-throughput manner.

Genome editing

The prerequisite of any project involving genome engineering, irrespective of its scale, was the development of methods capable of introducing precise and planned changes in

Submitted Apr 23, 2015; Accepted July 12, 2015
*Corresponding author. E-mail: feher.tamas@brc.mta.hu

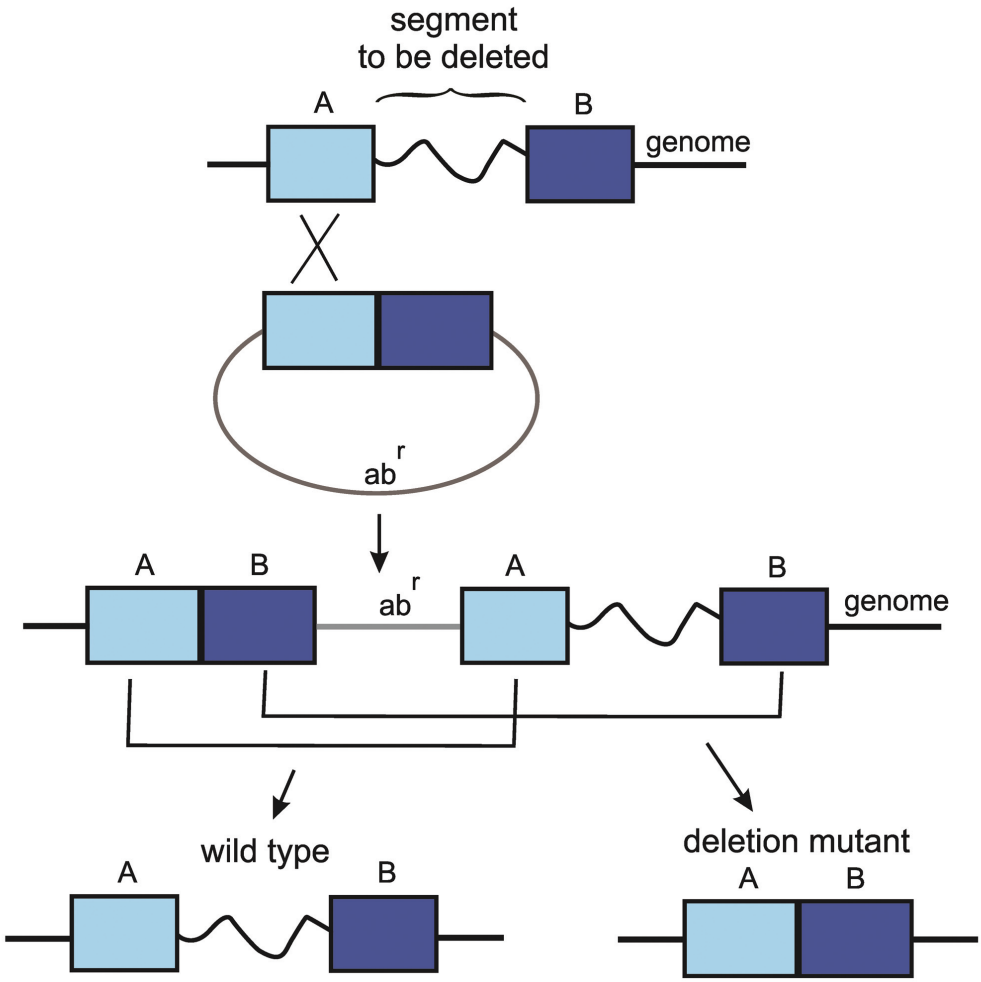

Figure 1. Suicide-plasmid mediated genome editing. Blue boxes represent regions of homology. ab^r: antibiotic resistance.

the genome of a bacterial cell. Therefore, before providing examples of strains engineered for biomedical studies or applications, we must give a brief introduction on the techniques available for such purposes.

Suicide plasmids

The earliest and most classical approach relies on conditionally replicative plasmids (Ruvkun and Ausubel 1981). These plasmids, also referred to as suicide plasmids, cannot replicate if the conditions change from permissive to non-permissive. In practice, this can mean a change in temperature (Hamilton et al. 1989), a change in the propagating host cell (Gutterson and Koshland 1983), or the loss of a protein factor required for replication (Miller and Mekalanos 1988). These plasmids, like most others, carry a selection marker (usually an antibiotic resistance gene). Non-permissive conditions cause the loss of the plasmid from the majority of the cells, leading to their death, or cessation of growth during antibiotic selection. A small subpopulation can survive, however, due to the integration of the plasmid into the genome. The process of integration is usually homologous recombination, mediated by the RecA recombinase of the cell. The locus of integration can be controlled by cloning a certain genomic segment into the plasmid, thereby providing a substrate for the homologous recombination (Fig. 1). The cloned genomic fragment is usually a mutated variant of the genomic allele, carrying an insertion, deletion or sequence alteration within its central part. In the simplest case, the integration itself can provide the experimenter a sufficient genomic modification, since it can mean the stable genomic propagation of a mutated gene, and all other genes originally present on the plasmid. However, this scenario results in the duplication of the targeted genomic segment, as well as the retaining of the antibiotic resistance gene within the genome, both of which can limit downstream experiments. Therefore, the excision

of the integrated plasmid is usually necessary. This process, referred to as the resolution of the cointegrant is an intramolecular homologous recombination, taking place between the duplicated segments (Fig. 1). The result can be either the wild type allele or the mutated allele remaining in the genome, depending respectively on whether the second recombination takes place in the same end or opposite end of the integrated genomic segment, relative to the site of integration (Fig. 1). The alternative outcomes of the resolution can be discriminated by colony-PCR, phenotypic tests, genome hybridization or sequencing. The resolution process itself is a rare event, but it can be selected for if a counterselective marker (*e.g.*, *sacB*) is present on the integrated plasmid (Blomfield et al. 1991). Better yet, resolution can be both mediated and selected for by either one of two elegant techniques: i) re-activating the origin of replication of the plasmid (Biswas et al. 1993), or ii) cleaving the cointegrant *in vivo* with a homing endonuclease, and thereby facilitating the second recombination step by providing free DNA ends (Pósfai et al. 1999). Importantly, suicide plasmid-mediated genome editing can be used to generate scarless genomic modifications, meaning that no exogenous sequence is left behind on the chromosome, besides the planned changes. Its drawback of requiring the cloning of relatively long (several hundreds of bps) stretches of DNA is counterweighted by the fact that it is freely available in the public domain.

Suicide plasmids have been used in complex schemes (*e.g.*, facilitating genome sequencing; Wild et al. 1996), but are mostly used to create simple knockout strains of pathogens. This is usually done with the aim of verifying the function of putative virulence genes, but hypovirulent strains applicable as vaccines can also be engineered this way. For example, the 13 kbp-long multi-drug resistance locus of *Salmonella enterica* serovar Typhimurium was deleted using a temperature sensitive suicide plasmid (Sahu et al. 2013). The deletion strain displayed decreased colonization and proliferation capabilities in *Caenorhabditis elegans*, as well as an altered immune response of the infected host. In another work, the genes encoding the lethal toxin and the edema toxin of *Bacillus anthracis* were deleted separately and in combination to study their roles in disrupting a human brain microvascular endothelial cell monolayer, or during the *in vivo* infection of mice (Ebrahimi et al. 2011). Interestingly, the two individual deletion mutants performed in an opposite fashion in the two virulence tests. A further project applying suicide plasmids to study virulence factors was the work of Horzempa et al. (2010), who deleted the *pyrF* gene of *Francisella tularensis* to obtain pyrimidine auxotrophic cells. These lost their ability to replicate in primary human macrophages, but surprisingly retained virulence during infection of chicken embryos and in the murine model of pneumonic tularemia. An example of vaccine development using suicide plasmids was published by Lee and co-workers (2007). In that work, the intranasal vaccination of mice with the *rpoS*, *phoP* double mutants *of Salmonella enterica* serovar Typhi was shown to provide effective immunity against the wild type strain.

Recombineering

The second widely used strategy, nowadays referred to as "recombineering" integrates linear DNA into the chromosome via a double homologous recombination process (Fig. 2) (Murphy 1998). Recombineering displays several important advantages over suicide plasmid-based methods: it does not require cloning, and the terminal homolog regions of the DNA fragment are short enough to be accommodated on primers. In several fungal and bacterial species, *e.g.*, *Saccharomyces cerevisiae* (Baudin et al. 1993)*, Bacillus subtilis* (Fabret et al. 2002) or *Sulfolobus acidocaldarius* (Grogan and Stengel 2008) the effectiveness of the housekeeping recombinase machinery allows the process to efficiently take place in wild type cells. In *Escherichia coli*, however, the expression of phage recombinases is usually necessary, either by activating the *Rec*ET operon of the endogenous Rac prophage (Kusano et al. 1994) or by transforming the *Red*αβγ genes of phage lambda (Murphy 1998). Analogously, expression of recombination-protein gp61 of mycobacteriophage Che9c (van Kessel and Hatfull 2007) has been exploited for genome editing in Mycobacteria. Phage-derived recombinase proteins, comprising an endonuclease and/or an ssDNA-binding protein have also been identified in *Vibrio cholerae* (Chen et al. 2011) and *Clostridium perfringens* (Dong et al. 2014), opening the possibility of applying recombineering for these species as well.

The classical scheme of recombineering applies double stranded linear DNA in a two step procedure, similarly to that of suicide plasmids. During the first step, an antibiotic is used to select for cells carrying the cointegrate. This step alone is sufficient to disrupt or completely delete a chromosomal gene (referred to as "knock-out"), or to integrate a gene (known as "transgene knock-in"). In most cases, however, unwanted exogenous sequences are removed in a second step, using either one of three methods. As a first alternative, site specific recombinases, *e.g.*, FLP, Cre or Vika can be used to facilitate recombination between two short target sites (FRT, *loxP* or *vox,* respectively) (Broach and Hicks 1980; Sternberg and Hamilton 1981), thereby removing all sequences in-between (Fig. 2A). This method is simple and efficient, but a recombinase target site is retained in the genome, which can interfere with downstream experiments. The second alternative is to apply a second round of recombineering using a DNA fragment that consists solely of the scarless joint (Fig. 2B). This usually requires inclusion of a counterselective marker on the initially inserted segment. The third technique for removing the unwanted section of the insert is to use three homology boxes, as shown on Figure 2C. The two terminal boxes war-

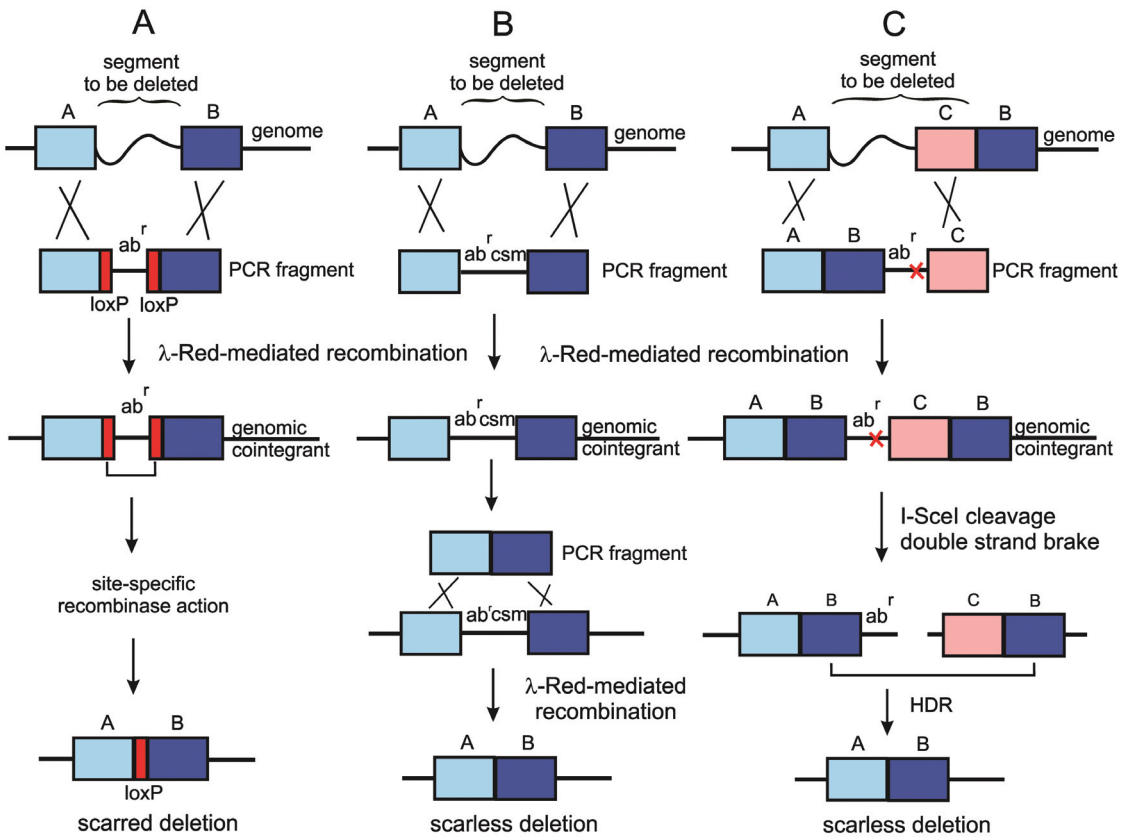

Figure 2. Recombineering. Three alternative strategies to engineer markerless genomic deletions are shown. Blue and pink boxes represent regions of homology, red boxes depict targets of a site-specific recombinase. ab^r: antibiotic resistance; csm: counterselective marker; HDR: homology-directed recombination.

rant integration into the genome, while a third, internal box homologous to a nearby genomic segment allows a second round of recombination, eliminating all sequences between this box and the nearby segment. The efficiency of this technique can be greatly enhanced by *in vivo* cleavage of the insert with a homing endonuclease (Jurica and Stoddard 1999) which provides free DNA ends for recombination and selection against insert-bearing cells. The two latter techniques are suitable for scarless genome engineering, since they leave no unwanted sequences behind.

Lately, an alternative scheme of recombineering is gaining ground, especially in high-throughput genome engineering. This method uses single stranded DNA, and only a single step procedure for genome editing (Ellis et al. 2001). Briefly, 40-80 bp long oligonucleotides corresponding to the lagging strand of DNA replication are electroporated into cells expressing the lambda Redβ recombinase protein. The efficiency of their recombination into the genome is further enhanced by eliminating the methyl-directed mismatch repair system of the host (Costantino and Court 2003), by using chemically modified nucleotides (Wang et al. 2011), or by overloading or eliminating host endonucleases (Sawitzke et al. 2011). Mutations can be introduced into the genome this way with frequencies between 0.1 and 50%, allowing simple screening by colony-PCR to find recombinants.

As demonstrated for suicide plasmids, recombineering is also primarily used to generate null-mutants, thereby aiding the identification of the function of the deleted gene. In a study describing a novel inhibitor of pyocianin production and biofilm formation of *Pseudomonas aeruginosa*, the putative target of the inhibitor was validated using null mutants of quorum-sensing receptors (O'Loughlin et al. 2013). Recombineering in *M. tuberculosis* applying the mycobacteriophage gp61 revealed that an asparagine transporter is necessary for this pathogen to assimilate nitrogen, as well as to deal with acid stress within the phagosome (Gouzy et al. 2014). Mutants lacking this transporter displayed attenuated growth in murine macrophages and infected adult mice. Removal of virulence factors to generate attenuated live vaccine strains was also demonstrated with λ-Red recombineering: deletion of the *virG* gene of *Shigella flexneri* led to a decreased invasion of epithelial cell monolayers as well as a lack of inflammation

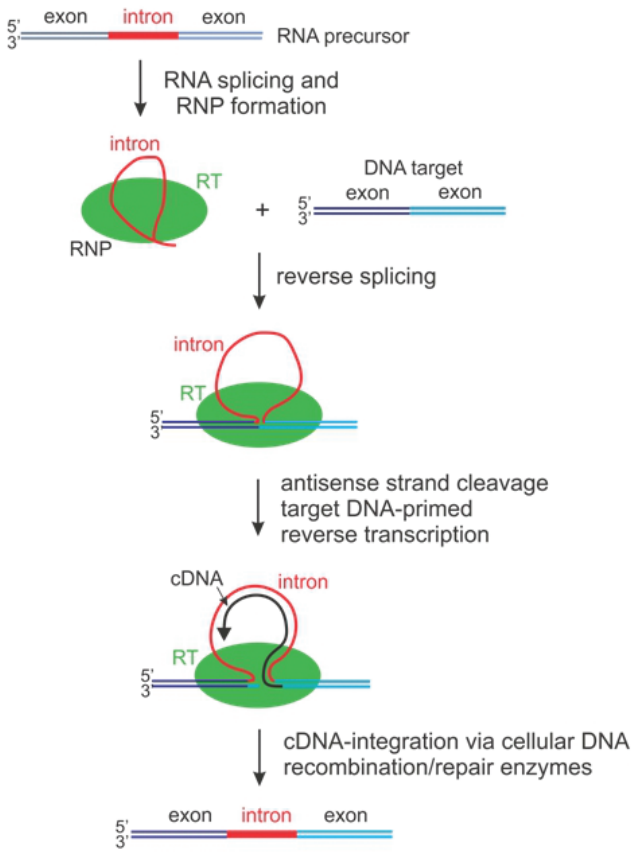

Figure 3. Retrohoming of a group II intron (Enyeart et al. 2013). RT: reverse transcriptase; RNP: ribonucleoprotein

upon challenging the conjunctiva of guinea pigs (Ranallo et al. 2006). Recombineering is also capable of editing episomal DNA, including high-copy plasmids (Thomason et al. 2007), bacterial artificial chromosomes (Muyrers et al. 1999), and bacteriophages. Phage genome editing, earlier done exclusively in the lysogenic state (Oppenheim et al. 2004) has lately been demonstrated to be possible in the lytic state as well, both with λ-Red (Fehér et al. 2012a) and with Che9c gp61 recombinase proteins (Marinelli et al. 2008). The use of engineered phages in curing model diseases is a promising field of synthetic biology (see below).

Group II introns

Less frequently used, but important tools of targeted genome-editing are the group II introns, also called targetrons (Lambowitz and Zimmerly 2004). The natural forms of these elements can increase their copy numbers in the genome of their host by a mechanism called retrohoming. During this process, the RNA transcribed from these elements goes through maturation, and forms a lariat structure (Fig. 3). This lariat can attack and insert into double-stranded DNA (dsDNA) by "reverse splicing", forming a single-stranded RNA insertion. This intermediate form is reverse transcribed, and subsequently repaired to form a dsDNA insert by the endogenous DNA repair mechanisms of the host cell. The RNA maturation, DNA cleavage and reverse transcription are all catalyzed by a single enzyme called the intron encoded protein (IEP). Using a computational tool (Perutka et al. 2004), the insertion process can be specifically retargeted by modifying the segments of the intron responsible for target binding. The biggest advantage of targetrons is their applicability in cells that are recalcitrant to homologous recombination, thereby providing the user an unusually broad host range genome–engineering tool (Rodriguez et al. 2009).

The most straightforward application of group II introns is to interrupt genes, similarly to their natural mode of action. They can be used to introduce cleavage sites into the chromosome, thereby allowing the generation of DSBs followed by their correction by homologous recombination of a co-transformed linear DNA fragment. Targetrons are also capable of inserting cargo genes (Frazier et al. 2003), or target sites of site-specific recombinases into the genome. With properly integrated recombinase target sites, one can engineer deletions, inversions or cut-and-paste translocations of chromosomal segments upon the cellular expression of the corresponding recombinase (Enyeart et al. 2013). The insertion frequency of *the Lactococcus lactis* Ll.LtrB group II intron is in the range of 1% in *E. coli*, (Lambowitz and Zimmerly 2004) meaning that PCR-based screening can be sufficient to find cointegrants even if using systems that lack a selection marker. If certain experimental setups or hosts require elevating the rate of insertion, one can apply selection when including a retrotransposition-activated selectable marker (RAM) in the intron. The first RAM was a Trimethoprim resistance gene, interrupted by a group I intron that goes through splicing during the retrotransposition process, thereby allowing the selection of co-integrants with nearly 100% specificity (Zhong et al. 2003).

Targetrons have been applied to engineer the genes of numerous gram-positive and gram-negative bacteria. To mention medically relevant examples, two putative replication origins of plasmid pXO1 from *Bacillus anthracis* have been disrupted one-by-one to validate their independent functionality, and measure the resulting plasmid copy numbers (Akhtar and Khan 2012). In another instance, the *mtlA* gene of *Vibrio cholerae* was knocked out, verifying its function in mannitol and sorbitol uptake (Kumar et al. 2010). A two-component regulator of *Yersinia pseudotuberculosis*, responsible for cold-tolerance was also inactivated using this technique (Palonen et al. 2011). Besides engineering knock-out strains for gene functionality studies, group II introns have also been used to construct transgenic bacterial strains for the purpose of vaccination. For example, recombinant *Clostridium perfringens*

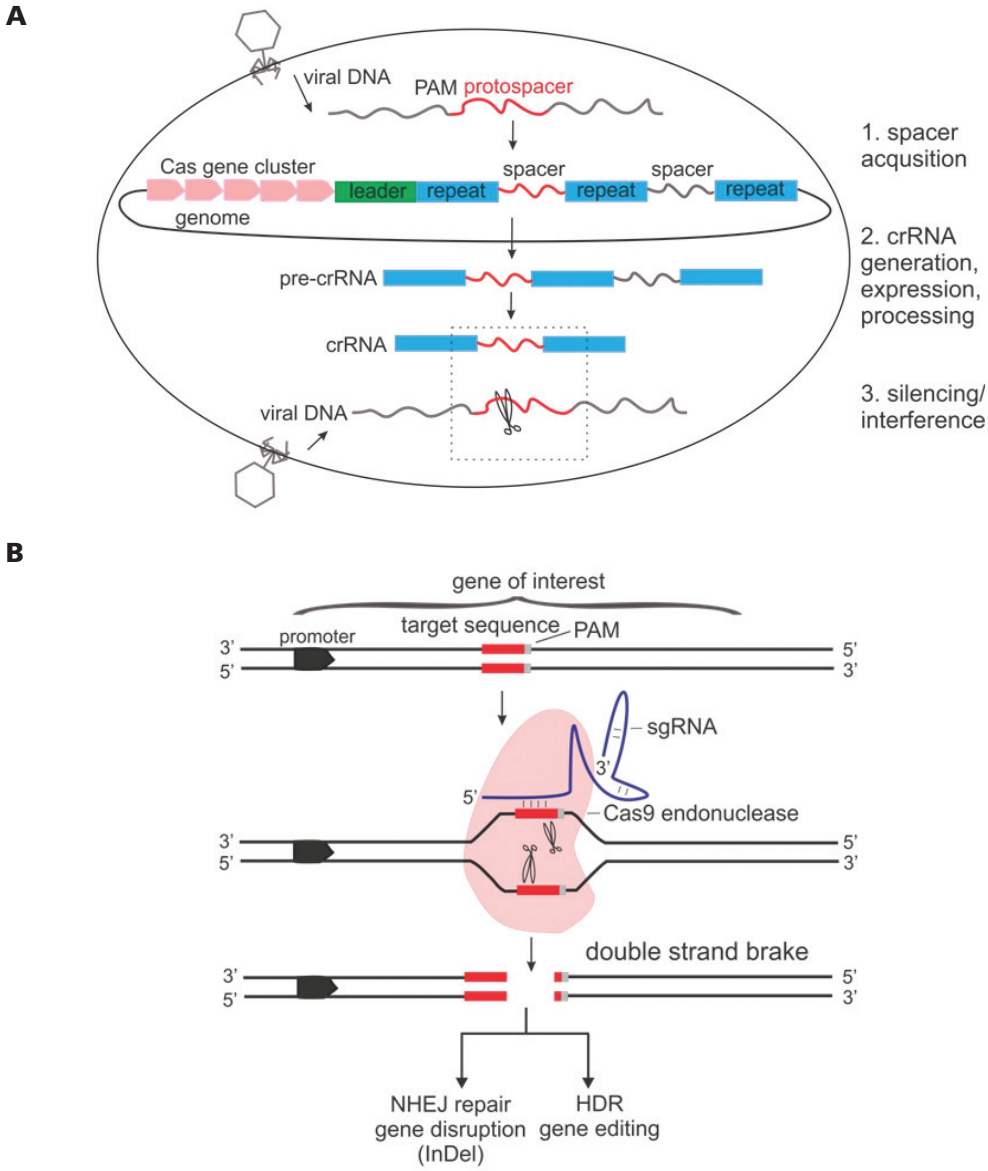

Figure 4. (A) Structure and function of a CRISPR array. (B) Utility of the Cas9 protein as an RNA-guided nuclease. The specificity of the cleavage is warranted by the base-pairing of the sgRNA and the target DNA. PAM: protospacer adjacent motif; NHEJ: non-homologous end joining; HDR: homology-directed recombination; InDel: insertions or deletions.

cells are potentially useful vehicles for oral delivery of antigens to the gut-associated lymphatic tissue (GALT). However, plasmid-borne expression of the antigen necessitates the use of resistance markers for plasmid maintenance, which would raise the risk of iatrogenic transmission of antibiotic resistance. Another hazard connected to clostridia is their potential to produce various toxins. To solve these two problems in one step, Chen et al. (2007) used a targetron to knock-in the p27 gene of the Simian Immunodefiency Virus into the theta toxin gene (*pfoA*) of the *C. perfringens* genome, thereby creating a safer strain for vaccination against the viral antigen. Another example of engineering an attenuated vaccine strain was the deletion of the 15 kbp-long pathogenicity island 1 of *Staphylococcus aureus* using two flanking targetrons and the Cre/lox site specific-recombinases (Enyeart et al. 2013).

The CRISPR/Cas system

Recently, a novel technique relying on the bacterial CRISPR/Cas system opened an entirely new chapter in genome edit-

ing, and most probably in general molecular genetics as well. The activity of CRISPR/Cas, in its wild type form is often referred to as the adaptive immune response of bacterial cells (Horvath and Barrangou 2010; Jinek et al. 2012). Its presence has been demonstrated in 45% of bacterial and 83% of archeal genomes (Sampson and Weiss 2014). The name CRISPR, which stands for clustered regularly interspaced palindromic repeats, originates from its genetic organization, shown on Fig. 4A. Key elements of the CRISPR loci are the direct repeats which alternate with variable spacers, remnants of captured exogenous plasmid and phage sequences. The spacers are responsible for specific recognition of their maternal DNA sequences upon their re-entry into the cell. The long primary transcript of the CRISPR locus (the pre-crRNA) is processed into short mature CRISPR RNAs (crRNAs) that consist of a single repeat and a spacer, with the latter targeting potential invader sequences. Importantly, the targeting process is a Watson and Crick base pairing, which leads to the cleavage of the target DNA by the Cas enzymes.

The Cas proteins make up a heterogeneous family harboring domains reminiscent of endonucleases, helicases, polymerases, and polynucleotide-binding proteins. Cas genes usually lie in the proximity of the repeat-structure, but are transcribed separately from the pre-crRNA. To date, three classes of Cas have been described, that are distinct in their molecular mechanisms. Type I and III utilize a complex of Cas proteins (called CASCADE) both for RNA maturation and for target DNA cleavage. In the case of the type II system, a single enzyme, Cas9 is sufficient to disrupt the target, which explains why most applications use this system for gene editing. It is important to note, that for type II CRISPR/Cas, the trans-activator CRISPR RNA (tracrRNA), complementary to the repeat of the CRISPR array, is required both for RNA maturation and the activation of Cas9. In practical applications using Cas9 as an RNA-guided endonuclease, the natural setup is further simplified by using a single guide RNA molecule (sgRNA) that unites the roles of the crRNA and the tracrRNA in target recognition and Cas9 activation, respectively (Fig. 4B). Importantly, for type I and II CRISPR/Cas, the targeted sequences must include a protospacer adjacent motif (PAM), which slightly limits the freedom of guiding the cleavage process. Today, several websites are available that aid the design of sgRNAs (http://www.e-crisp.org/E-CRISP/designcrispr.html; http://zifit.partners.org/ZiFiT/) (Sander et al. 2010; Heigwer et al. 2014).

The rapid expansion of CRISPR/Cas-based genome engineering projects derives from its three major merits. The first is its adaptability to multiple domains of life. For example, the type II system of *Streptococcus pyogenes* has been applied in species as diverse as *E. coli* (Jiang et al. 2013), yeast (DiCarlo et al. 2013), *C. elegans* (Friedland et al. 2013), *Drosophila* (Gratz et al. 2013), *Danio rerio* (Hwang et al. 2013), rice (Feng et al. 2013) and human cell lines (Cho et al. 2013). The second advantage is the simplicity of design. Earlier programmable nucleases like zinc finger nucleases, transcription activator-like effector nucleases (TALENS) or engineered meganucleases required protein design, which significantly increased the time, effort and financial requirements of the engineering project (for a review of these systems, see (Sun et al. 2012)). Since the CRISPR/Cas system relies solely on base-pairing of the guide RNA with the target, the design process is incomparably simpler. The third merit is the flexibility of the system, which gave way to multiple derivative techniques, discussed below.

The basic application of CRISPR/Cas is introducing a double-stranded break (DSB) into the target DNA. In eukaryotes, this is usually sufficient to inactivate the gene, since non-homologous end joining, the primary system repairing DSBs often leads to a frameshift mutation (Sun et al. 2012). For more precise and predictable genome engineering, a small linear DNA fragment, referred to as the editing template is provided along the cleavage process, which bridges the free DNA ends using homology-directed repair. With proper design of the editing template, one can engineer arbitrary genetic alterations, including gene knock-outs and knock-ins. Mutants of the Cas9 protein that nick the target DNA instead of cleaving it, have been applied to increase target specificity: by introducing nicks to opposite DNA strands at two adjacent sites, one can effectively double the length of the target identified during complete cleavage (Ran et al. 2013). This strategy can increase specificity, similarly to the classic trick of extending restriction endonuclease target sites (Pósfai and Szybalski 1988). Cas proteins that completely lack endonuclease activity, called dead Cas9 (dCas9) have been applied to alter gene expression by either one of two mechanisms: they can either inhibit transcription initiation via promoter-binding, or can be fused to a transcription activator or repressor and promote expression up- or down-regulation, respectively. Another application of the specific targeting function is using dCas-GFP fusions for fluorescent visualization of certain genomic segments *in vivo*.

The amazing popularity that CRISPR/Cas has gathered in the past two years is mostly attributable to its applicability in mammalian cells. Not surprisingly, due to the alternative possibilities available in prokaryotes, the number of publications reporting its application grows at a much slower rate. The possibilities are nevertheless given, as exemplified by the following proof-of-concept studies. For example, the type II CRISPR/Cas of various *Streptococcus* species has been successfully applied to edit the genomes of *S. pneumoniae*, *E. coli*, *Lactobacillus reuteri* and *Streptomyces coelicolor* (Jiang et al. 2013; Oh and van Pijkeren 2014; Huang et al. 2015). The DSB introduced by Cas has been shown to both facilitate the recombination of the editing template, and provide a selection against the wild-type genotype (Jiang et al. 2013). In *E. coli* and *L. reuteri*, the function of the editing templates

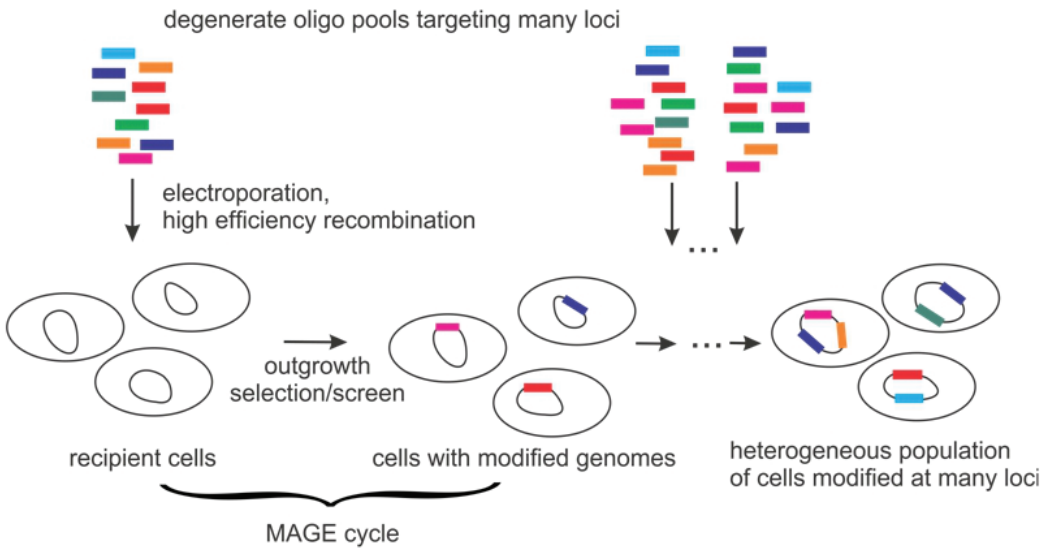

Figure 5. Multiplex Automated Genome Engineering. Boxes of different color represent oligonucleotides targeting different genomic loci, or oligonucleotides targeting the same locus but carrying different mutations.

can be fulfilled by oligonucleotides if the recombination is mediated by phage recombinases, thereby combining recombineering with the CRISPR/Cas system. If no bridging DNA is provided, the DSB causes cell death in most bacteria, and has been demonstrated to be applicable in removing bacterial strains from a mixed population based on sequence content (Gomaa et al. 2014). This strategy can be used to generate "smart antibiotics", targeting only those bacteria that present a danger to the host (see below). For gene expression down-regulation, also called CRISPR-interference (CRISPRi), both type I and type II CRISPR/Cas has been applied in *E. coli* (Bikard et al. 2013; Luo et al. 2015). The engineering of bacteriophages has also been demonstrated with these two types of CRISPR/Cas (Kiro et al. 2014; Martel and Moineau 2014). Perhaps the largest therapeutic promise of CRISPR/Cas systems lies in the *in vivo* inactivation of proviral DNA inside eukaryotic cells (see below).

Large scale engineering of microbial genomes

As soon as techniques for markerless genome editing were available, the door was open for repeated engineering of bacterial or fungal chromosomes. Projects that have introduced multiple targeted changes into genomes are numerous. We will concentrate only on two special cases: genome reduction, and multiplex parallel (automated) genome engineering.

Genome reduction

The first publication that reported the reduction of a microbial genome was the work of Kolisnychenko et al. (2002), who applied recombineering to remove the 12 largest strain-specific islands of *E. coli* K-12. Since then, genome reduction has been applied in at least six bacterial (*E. coli, Bacillus subtilis, Corynebacterium glutamicum, Legionella pneumophila, Pseudomonas putida, Streptomyces avermitilis)* and three fungal *(Saccharomyces cerevisiae, Schizosaccharomyces pombe, Aspergillus oryzae)* species (for a recent review, see Fehér 2014). The motivation behind most projects is usually one of the two following factors: i) to answer fundamental questions, *e.g.*, seeking the minimal gene set required for life, investigating the evolutionary role of mobile elements (Fehér et al. 2012b) or ii) optimization of the strain for biotechnological applications, *e.g.*, gene cloning or protein overproduction. In the latter case, unwanted cellular components are removed with the aim of increasing cellular predictability and stability, or redirecting energy and metabolite resources towards the production of useful compounds (Umenhoffer et al. 2010). An example that applies genome reduction for improved production of pharmaceuticals is the work of Komatsu and co-workers (2010). In their work, they deleted the 1.5 Mbp-long left subtelomeric region of the linear *S. avermitilis* chromosome. This eliminated the production of the endogenous secondary metabolites avermectin and filipin, and two further deletions removed the entire gene set of oligomycin biosynthesis. As a result, the reduced-genome strain, when transformed with the appropriate genetic constructs displayed increased production yields of certain pharmaceuticals, including aminoglycosides, polyketides, non-ribosomal peptides and terpenes. In fact, the engineered strain produced streptomycin and cephamycin with higher yields than the *Streptomyces* strains that the heterologous pathways originated from.

Highly parallel genome engineering

In 2009, the group of George Church introduced the quantitative upscaling of the recombineering process (Wang et al. 2009). Multiplex Automated Genome Engineering (MAGE) employs the λ-Red recombinase and a pool of oligonucleotides to target multiple genomic loci, simultaneously (Fig. 5). No direct selection is applied for individual recombination events; one can however employ a general selection pressure to direct the phenotype in the direction of choice. The oligonucleotides can be designed to introduce defined changes (insertions, deletions or base exchanges), but can also carry degenerate segments to allow randomization of the targeted region. This way the experimenter can implement directed evolution where the increased mutation rate is limited to a few loci of interest. The application of MAGE in repeated cycles allows generating a tremendous combinatorial library of mutants and the exploitation of genetic interactions where the contributions from multiple neutral mutations result in a fitness increase.

In a proof of principle experiment, MAGE was used to target 24 genes related to lycopene production on the *E. coli* genome. Apart from deleting some draining pathways, most of the oligos were designed to introduce changes in the ribosome-binding sites of genes involved in production, thereby altering their translation rate. This generated a combinatorial library of mutants, where the balanced overexpression of enzymes, resulting in maximal lycopene production, was selected relying on the color of the colonies (Wang et al. 2009). In a later work, the power of MAGE was combined with the specificity of transcription factors sensitive to the target compound. This allowed linking target production to cell fitness and automatic selection of the best producer strains from a vast library of mutants, thereby increasing naringenin and glucaric acid production rates 36- and 22-fold, respectively (Raman et al. 2014). One potential disadvantage of MAGE is the requirement of a mismatch-repair defective (*mutS*$^{-}$) host cell to potentiate the recombination process, which unintentionally increases the overall mutation rate of the genome 100 fold. Recently, the rate of such off-target mutations has been decreased by applying a system that inactivates *mutS* only temporarily, thereby warranting the stable maintenance of engineered strains (Nyerges et al. 2014).

Genome synthesis

Genome synthesis is the *de novo* assembly of DNA molecules corresponding to the genome of existing or planned microorganisms. The basic building blocks of this process are the oligonucleotides, chemically synthesized ssDNA molecules usually falling into the size range of 30 to 150 bases. The toolbox available for genome synthesis is quite vast, due to the fact that stitching together pieces of DNA has been carried out for decades within the bounds of genetic cloning. Although some of these methods are very elegant and would deserve a detailed description, discussing each of them would exceed the limits of this publication. We therefore concentrate only on the ones that have already been applied in genome synthesis, and direct the reader to earlier reviews to obtain a complete picture on all the possibilities of DNA assembly (Ellis et al. 2011; Cobb et al. 2013).

Assembling oligonucleotides to longer stretches of dsDNA, referred to as gene synthesis is frequently outsourced to commercial providers today. There are two basic *in vitro* methods to carry out this task. The first one relies on overlapping phosphorylated oligonucleotides that completely span both strands of the target DNA. These are stitched together in the course of ligase chain reaction (LCR), a cyclic process of annealing and ligation (Grundstrom et al. 1985). The second method uses partially overlapping oligonucleotides that only partially cover the two strands. The gaps in between two neighboring oligos are filled up by Taq polymerase in overlap-extension reactions, using the complementary strand as a template. After numerous cycles of this reaction, called Polymerase Cycling Assembly, a significant portion of the products will be the nick-free full length target DNA molecule (Stemmer et al. 1995). The first method is more expensive, the second one is more error-prone, and both require sequence-verification of the final product. A combination of these two strategies has been successfully applied to assemble the 5.4 kbp genome of the PhiX174 phage from 42 bp oligonucleotides in a time frame of two weeks (Smith et al. 2003).

Several methods have been used to assemble gene-sized fragments into genomes. The first and most fundamental one is restriction-ligation. Despite its numerous drawbacks (slowness, size limitation of fragments, limited re-use of sites, sequence scars remaining at the joints), it was applied in the first assembly of a synthetic polyovirus (Cello et al. 2002). Although several *in vitro* methods have been developed to circumvent these disadvantages, only one of them has been applied to date in genome synthesis. This method, nowadays referred to as Gibson assembly, is an *in vitro* recombination of overlapping dsDNA fragments. In the most popular version of the Gibson technique, the T5 exonuclease is used to chew-back the 5' ends of the dsDNA molecules. The emerging 3' overhangs are extended by the Taq polymerase after annealing and are ligated by the Taq ligase, all in an isothermal reaction at 50 °C (Gibson et al. 2009). This method has been used to assemble the 16.3 kbp-long circular genome of murine mitochondria in multiple hierarchical rounds, starting from 600 overlapping oligonucleotides (Gibson et al. 2010a). It has also been demonstrated to be capable of assembling the 590 kbp *Mycoplasma genitalium* genome from four pieces (Gibson et al. 2009).

Another major strategy to assemble genome-sized DNA is to use the remarkable, *in vivo* recombination capacity of

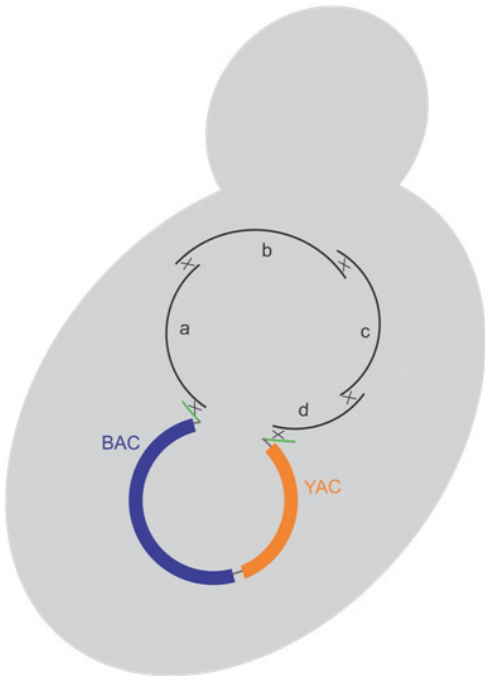

Figure 6. Transformation Associated Cloning in yeast (Gibson et al. 2008b). Thin lines represent the DNA to be assembled and cloned. Blue and orange boxes represent bacterial and yeast replication origins, respectively. BAC: bacterial artificial chromosome vector backbone, YAC: yeast artificial chromosome vector backbone.

S. cerevisiae. This method, called transformation associated recombination (TAR) cloning only requires having 60 bp overlaps among the transformed dsDNA segments and including a yeast replication origin to allow forming a circular yeast plasmid that comprises the entire synthetic genome (Fig. 6) (Gibson et al. 2008b). The power of the method has been demonstrated in the assembly of the 590 kbp-long *M. genitalium* genome from 25 pieces (Gibson et al. 2008a), or the construction of the 1.08 Mbp-long *M. mycoides* genome in multiple hierarchical rounds, starting from synthetic DNA fragments of 1 kbp (Gibson et al. 2010b).

If a bacterial genome is assembled *in vitro*, or inside a eukaryotic cell (as in the examples above), a special technique is required to insert it into a bacterial cell to be "booted". This special form of transformation, called whole genome transplantation (Lartigue et al. 2007) was used to insert a synthetic *M. mycoides* genome into *M. capricolicum* cells and obtain *M. mycoides* offspring controlled by the synthetic genome. The obtained cells were defined as synthetic, and are hallmarks of bottom-up engineering of bacterial genomes. However, genome-scale cellular models do not currently permit the assembly of genomes with a gene content significantly differing from those of wild type cells. For trial-and-error type of genome engineering, editing existing genomes is incomparably simpler and less expensive. The major value of the developed methodologies (TAR cloning and whole genome transplantation) seems to be the ability to clone the genomes of species for which no genome engineering techniques are available. These genomes could be edited in yeast, and then re-transplanted into wild-type cells to obtain engineered offspring (Fig. 7) (Lartigue et al. 2009).

Although the power of genome assembly is readily apparent, its affordability is yet to be seen. Currently, only one project has been published that assembled a microbial genome for biomedical purposes. The workgroup that synthesized the 7.5 kbp-long polyovirus genome discovered that introducing novel restriction sites in coding regions as silent mutations, necessary for the assembly had a severe deleterious effect on the *in vivo* neurovirulence of the virus (Cello et al. 2002). To turn this side effect into a benefit, the group re-synthesized the poliovirus genome, this time intentionally using rare codon-pairs throughout the capsid-encoding region. The result was a strain with perfect immunogenicity but 10 000-fold reduced efficacy in murine neurovirulence assays due to a decreased rate of translation. The synthetic virus elicited a protective immune response against wild-type poliovirus in a mouse model. Such viral strains could be ideal vaccine candidates, since their regaining of virulence would require a high number of reverse mutations and therefore has a negligible chance (Coleman et al. 2008).

Promoting health *via* microbial genome engineering

This section will showcase notorious examples of using engineered bacteria and bacteriophages for treating or preventing various diseases. Metabolic engineering projects, *i.e.* those involving the microbial production of drug precursors, neutraceuticals or other fine chemicals intended for medical use will not be discussed, for these represent a separate and concise field on their own (for a review, see Woolston et al. 2013). Nor will we discuss the use of wild type bacteria or bacteriophages as probiotics or as phage-therapy tools, respectively. We must note however, that the experience accumulated by the two latter disciplines provided a strong support for many of the following applications that use the genetically-engineered derivatives of the wild type agents.

Engineering live, attenuated bacterial strains by knocking out virulence genes for vaccination purposes is a long-proven strategy. An even more resourceful approach is to alter, or broaden the immunogenicity of a strain, thereby eliciting an immune response that is protective against further species. The first example of the latter dates back to 1993, when Connell and co-workers expressed the gp63 surface proteinase of *Leishmania* in the cytosol of *Mycobacterium bovis* BCG cells (Connell et al. 1993). Mycobacterial cells were chosen as the vaccine delivery vehicle for they had been proven to

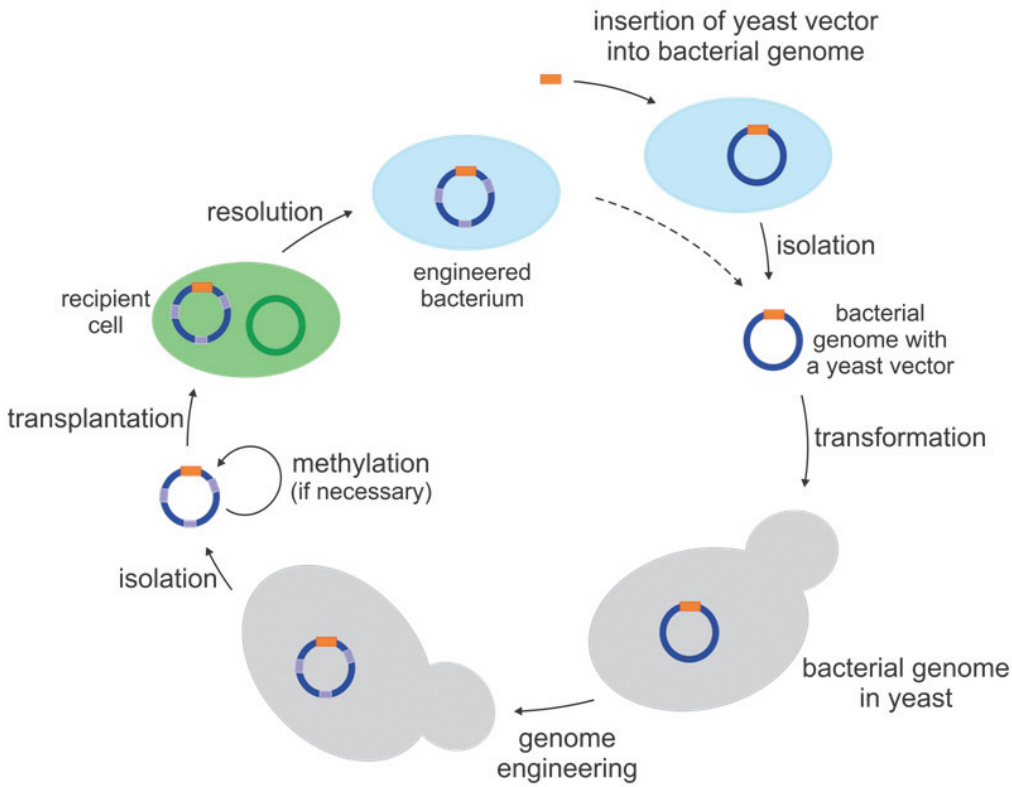

Figure 7. Cloning and engineering bacterial genomes in yeast (Lartigue et al. 2009). After the engineering is complete, the genomes are purified and transplanted into wild type bacterial cells.

be an effective adjuvant for the induction of a protective T_{H1}-mediated cellular immune response to leishmaniasis. Intravenous vaccination of mice with engineered BCG cells led to a marked decrease of the diameter or the rate of progression of dermal lesions caused by a *Leishmania mexicana* challenge, although the extent of the response was dependent on the murine strain.

The use of virulent bacteriophages to lyse pathogenic bacteria has a century-long history (Pennazio 2006). In the western world, the major application of this technology today is limited to the food industry and veterinary sciences (Atterbury 2009). Applying this process for humans *in vivo*, known as phage therapy, has had multiple eras of renaissance, but is today almost exclusively confined to countries formerly belonging to the USSR. However, a second generation of phage therapies seems to be emerging lately, which could provide yet another chance for this approach to gain ground in the west. This novel strategy applies genetically-engineered bacteriophages, and regards phages as mere delivery vehicles capable of inserting genetic constructs into bacteria for the purpose of their reprogramming. In a simple example, (Lu and Collins 2007) the T7 phage was engineered to express the *dspB* gene upon infection of the target cells. *dspB* encodes the dispersin enzyme, which is capable of degrading β-1,6-N-acetyl D-glucosamine, the major extracellular polysaccharide component of biofilms formed by *E. coli*. Despite the fact that phages can only access bacteria residing at the surface of the biofilm, lysis of the infected cells will release dispersin to initiate the degradation of the polysaccharide matrix (Fig. 8). This allows the released phage to re-infect deeper-lying cells of the biofilm, and ultimately to eradicate the targeted bacteria. In an *in vitro* test, the refactored phages were shown to decrease cell counts by 4.5 orders of magnitude, which meant a 100-fold increase in efficiency compared to the wild type phage.

In a more complex example, a non-lytic filamentous phage was used to influence a bacterial genetic circuit that is not targeted by current antibiotics (Lu and Collins 2009). This circuit was the SOS system, which can decrease cellular sensitivity to certain antibiotics by activating DNA-repair mechanisms in response to antibiotic-induced DNA-damage. In a proof-of concept experiment, phage infection generated lysogenic prophages expressing the *lexA3* allele inside the targeted cells, which is an uninducible and dominant mutant repressor of the SOS circuit (Fig. 9). As a result, the number of cells surviving ofloxacin treatment was reduced nearly

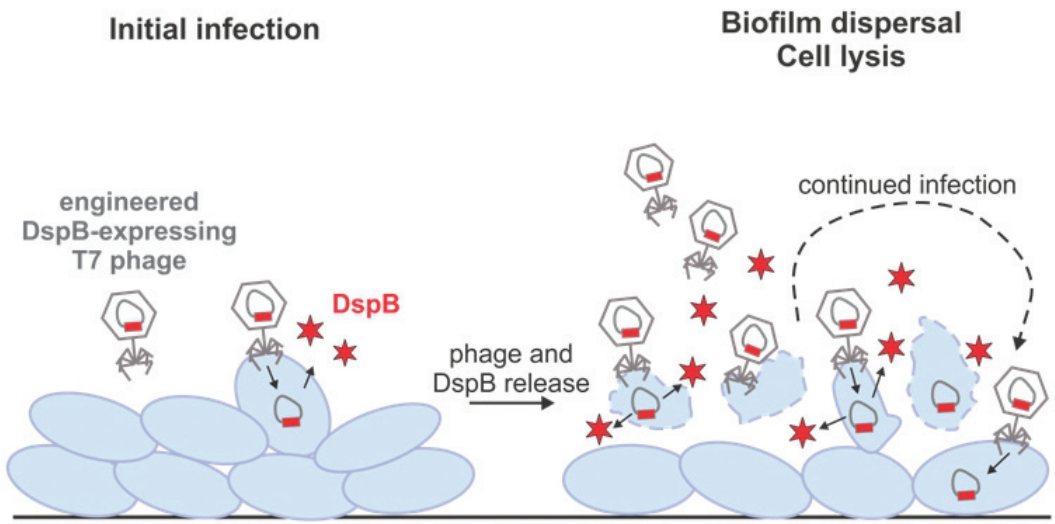

Figure 8. Biofilm dispersal using engineered bacteriophages (Lu and Collins 2007). Dispersin, the product of the phage-encoded *dspB* gene is responsible for degrading the extracellular polysaccharide matrix of the biofilm and providing access to deeper lying cells in repeated cycles of infection and lysis.

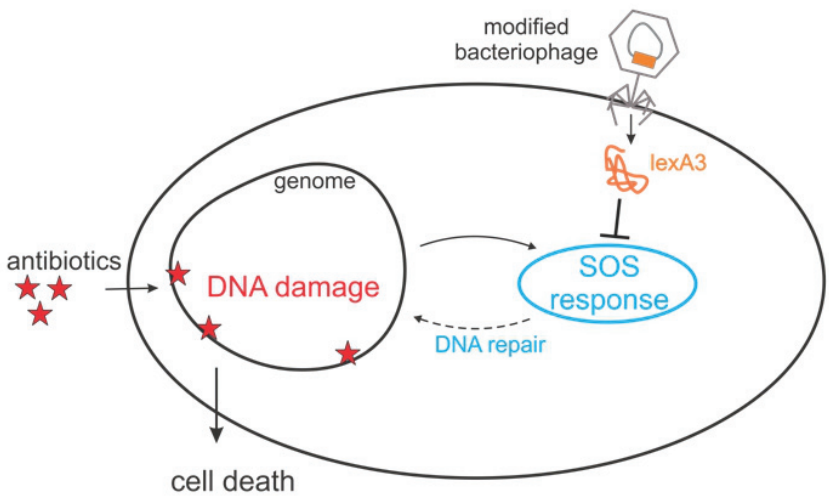

Figure 9. Increasing sensitivity to antibiotics *via* repression of the SOS response (Lu and Collins 2009). The product of the *lexA3* gene inserted into the bacterial cell by the transducing phage dominantly represses the SOS regulon, thereby avoiding the repair of the DNA damage caused by certain antibiotics.

10,000-fold *in vitro*. In an *in vivo* experiment, the survival rate of mice intraperitoneally injected with *E. coli* was elevated from 20% to 90% when the ofloxacin treatment was complemented with the engineered bacteriophage. In the same publication, the workgroup demonstrated the possibility of increasing antibiotic-susceptibility via three further mechanisms: i) expressing a porin (OmpF) that allows the uptake of an antibiotic; ii) overexpressing the repressor (SoxR) of an operon encoding anti-oxidative enzymes; and iii) expressing a master regulator (CsrA) to repress biofilm formation. In all examples, the phage treatment significantly facilitated the antibiotic effect *in vitro*, indicating the general applicability of this strategy.

The utility of phages as delivery vehicles has also been exploited in the development of "smart antibiotics". These are CRISPR/Cas systems that target bacterial genes which pose a danger to the infected organism, *e.g.*, antibiotic resistance or virulence genes. In one of the first applications, an RNA-guided nuclease specific for the kanamycin-resistance gene was introduced into *S. aureus* cells using a phagemid

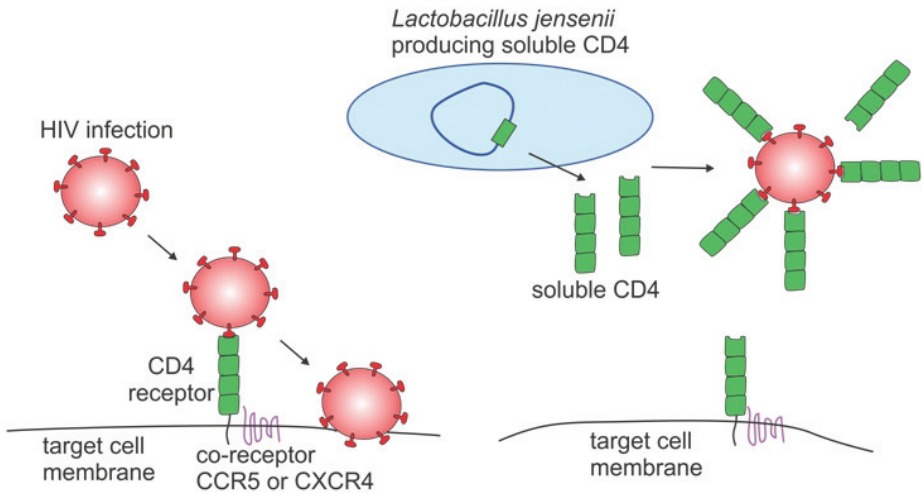

Figure 10. Inhibition of HIV binding by engineered *Lactobacilli* (Chang et al. 2003). The soluble CD4 molecules produced by modified *Lactobacillus jensenii* inhibit the docking of HIV with CD4 receptors residing in the membrane of target cells.

(Bikard et al. 2014). If the cells carried the resistance gene on the chromosome, their numbers could be reduced by up to 4 orders of magnitude due to the lethal effect of chromosomal cleavage. When the resistance gene was on a plasmid, 99.99% of the population became sensitive to the antibiotic upon phagemid treatment, without substantial changes in the cell number. In addition, the retained phagemid provided the host bacteria immunity to re-transformation with the resistance plasmid. In a parallel publication, conjugative plasmids were also demonstrated to be effective vectors for delivering RNA-guided nucleases into target cells (Citorik et al. 2014). In the same article, the survival curve of *Galleria mellonella* larvae infected with enterohaemorrhagic *E. coli* was shown to be significantly improved upon treatment with phagemids targeting a key virulence gene of the bacterium. Overall, both workgroups demonstrated the possibility of delicately restructuring mixed microbial populations, and nearly completely eliminating lines that carry potentially hazardous genes. Since there is no massive elimination of commensals, the remaining small fraction of targeted cells cannot go through a rapid expansion, as it often happens after incomplete antibiotic treatment. This warrants an extra level of safety for the future therapeutic use of RNA-guided nucleases.

Although this review is not intended to deal with the engineering of eukaryotic cells, the *in vivo* cleavage of proviral DNA using CRISPR/Cas should not be dismissed, for it stands on the borderline between targeted elimination of microbial parasites and gene therapy. The first such application was that of Ebina et al. (2013), who targeted the LTR region of the human immunodeficiency virus (HIV) provirus, which plays a central role in the regulation of proviral gene expression. Using GFP-encoding proviral DNA integrated into immortalized human cell lines, they demonstrated that RNA-guided cleavage of he LTRs disrupted the HIV-1 expression machinery, suppressed proviral re-activation and promoted proviral excision. A similar strategy was used to destruct intrahepatic persistent hepatitis B virus (HBV) in a mouse model (Lin et al. 2014). The CRISPR/Cas encoding plasmid was introduced into mouse hepatocytes *in vivo* using hydrodynamics-based transfection (Suda and Liu 2007), which led to the significant decrease of the serum levels of the viral surface antigen. Cleavage of the oncogenes of the human papilloma virus (HPV) provirus *in vivo* may also be a promising approach to inhibit tumorigenesis, as demonstrated in a mouse model (Zhen et al. 2014). An HPV-16 positive human cervical cancer cell line was transfected with the CRISPR/Cas plasmid targeting the E6 and E7 HPV oncogenes, as well as their promoter. The treated cells displayed reduced growth when transplanted into nude mice.

Modified bacterial cells may provide future possibilities to inhibit viral infection in the first place. In a proof-of-principle study, Chang et al. (2003) engineered a natural vaginal isolate of *Lactobacillus jensenii* to secrete a soluble form of CD4, the primary cell surface receptor responsible for HIV-docking and entry (Fig. 10). A significant inhibition of viral infectivity *in vitro* was attainable with a simple co-incubation of the engineered bacteria with the viral particles. This demonstrates that reinforcing the human microflora with engineered strains could prove to be a promising strategy for successful control of infections. In a similar study, the *E. coli* Nissle strain was engineered to secrete an inhibitor peptide of gp41, the HIV-protein responsible for membrane fusion and entry into target cells (Rao et al. 2005).

Engineered bacteria can also be used to prevent diseases

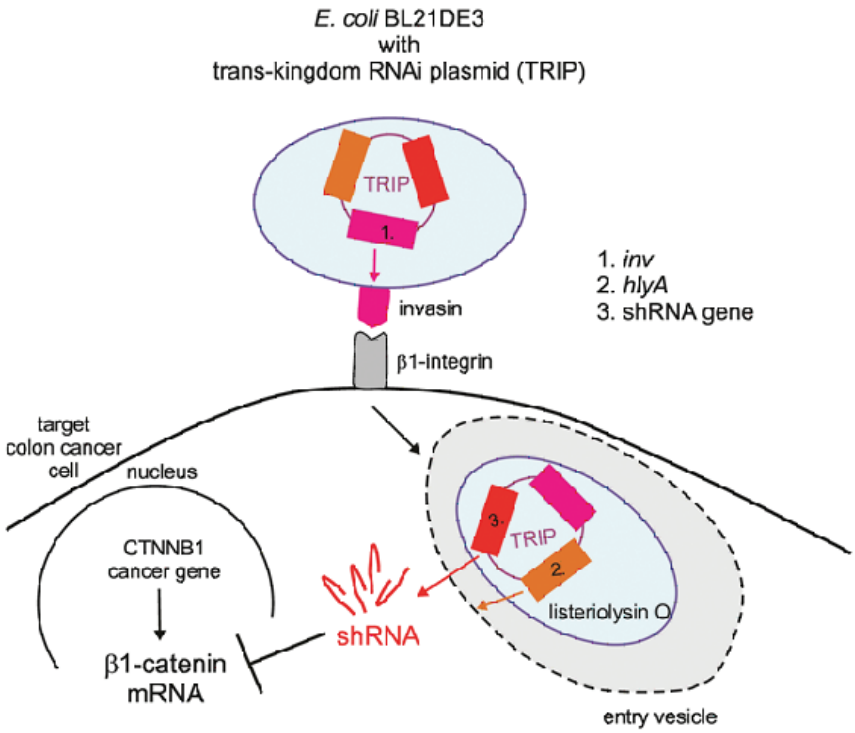

Figure 11. Trans-kingdom RNA interference (Xiang et al. 2006). *E. coli* cells engineered to be invasive secrete short-hairpin RNAs (shRNA) with the help of lysteriolysin O to knock down the target CTNNB1 gene encoding β-catenin.

by influencing pathogenic bacterial strains through intercellular signaling. For example, the growth and toxin-production of *Vibrio cholerae* is inhibited by the cholera autoinducer-1 (CAI-1) molecule produced by *V. cholerae* at high cell densities. An engineered *E. coli* strain that produces CAI-1 could therefore serve as a protective agent against gastrointestinal cholera infection. Indeed, pre-treating mice with a CAI-1 producer *E. coli* Nissle strain was shown to reduce intestinal cholera toxin quantities and *V. cholerae* cell counts by 80 and 69%, respectively, and improve host survival by up to 92% (Duan and March 2010). The application of refactored probiotic strains interfering with paracrine signaling therefore seems to be a useful strategy, and is by no means limited to treating or preventing infections. A summary of other engineered commensal strains secreting various factors to modulate human diseases is shown in Table 1.

One can apply bacteria to affect human cells not only through secreted biomolecules, but also by direct contact and entry. The anti-cancer effect of certain obligate or facultative anaerobic bacteria is long known, for such cells display a tendency to grow and become enriched in the poor oxygen conditions typical of malignant tumors (for a review, see Hoffman 2012). The development of precise genome engineering methods and their application in bacteria nevertheless opened brand new possibilities in this field of research as well. In a ground-breaking experiment, the *inv* gene of *Yersinia tuberculosis*, encoding invasin was shown to be sufficient to convert transformed *E. coli* into a microorganism able to invade

Table 1. Examples of engineered probiotic bacterial strains.

Strain	Modification	Effect	Reference
Lactococcus lactis	trefoil factors	wound healing, cytoprotection	Vandenbroucke et al. 2004, Caluwaerts et al. 2010
Lactococcus lactis	LcrV	anti-inflammation	Foligne et al. 2007
Lactococcus lactis	IL-10	anti-inflammation	Schotte et al. 2000, Steidler et al. 2000
Lactococcus lactis	HO-1	anti-inflammation	Pang et al. 2008, Pang et al. 2009
Lactococcus lactis	TNF-α nanobodies	anti-inflammation	Vandenbroucke et al. 2010
E. coli Nissle 1917	hEGFR	wound healing	Choi et al. 2012

cultured mammalian cells (Isberg et al. 1987). Controlling *inv* expression by the hypoxia-regulated *fdhF* promoter allows the construction of bacteria acquiring the invasive phenotype only in hypoxic conditions, thereby limiting their intracellular action to the cores of tumors (Anderson et al. 2006). Expressing the lysteriolysin O of *Lysteria monocytogenes* in such engulfed cells allows the release of their cellular components for transfer to the invaded host (Grillot-Courvalin et al. 1998). In a proof-of-concept study, *E. coli* cells expressing invasin and lysteriolysin O were used to transmit shRNA targeting β-catenin into the invaded cells (Fig. 11). Intravenous treatment of a mouse model of colon cancer with the engineered *E. coli* cells led to the significant decrease of the mRNA and protein levels of β-catenin within the tumor (Xiang et al. 2006). Although the therapeutic effect of modified *E. coli* in cancer therapy is yet to be seen, such experiments readily highlight the possibility and potential of using trans-kingdom RNA interference for reprogramming tumor cells.

Conclusive remarks

In the sections above, we tried to give a complete picture on the toolbox available to engineer the genomes of microbial strains. The examples highlighted, however are by no means concise, and represent only the tip of the iceberg concerning the application of engineered bacteria and phages for improving human health. By illuminating the available techniques and the key strategies followed, our primary aim was to raise the interest of the reader, and provide the basics necessary to initiate the next project in this rapidly extending field.

Acknowledgements

This work was supported by the TÁMOP-4.1.1.C-13/1/KONV-2014-0001 program entitled „Practice-oriented, student-friendly modernization of the biomedical education for strengthening the international competitiveness of the rural Hungarian universities". The authors thank Gabriella Balikó and György Pósfai for proofreading and helpful comments.

References

Akhtar P, Khan SA (2012) Two independent replicons can support replication of the anthrax toxin-encoding plasmid pXO1 of *Bacillus anthracis*. Plasmid 67:111-117.

Anderson JC, Clarke EJ, Arkin AP, Voigt CA (2006) Environmentally controlled invasion of cancer cells by engineered bacteria. J Mol Biol 355:619-627.

Atterbury RJ (2009) Bacteriophage biocontrol in animals and meat products. Microb Biotechnol 2:601-612.

Baudin A, Ozierkalogeropoulos O, Denouel A, Lacroute F, Cullin C (1993) A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. Nucleic Acids Res 21:3329-3330.

Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res 41:7429-7437.

Bikard D, Euler CW, Jiang W, Nussenzweig PM, Goldberg GW, Duportet X, Fischetti VA, Marraffini LA (2014) Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. Nat Biotechnol 32:1146-1150.

Biswas I, Gruss A, Ehrlich SD, Maguin E (1993) High-efficiency gene inactivation and replacement system for gram-positive bacteria. J Bacteriol 175:3628-3635.

Blomfield IC, Vaughn V, Rest RF, Eisenstein BI (1991) Allelic exchange in *Escherichia coli* using the *Bacillus subtilis sacB* gene and a temperature-sensitive pSC101 replicon. Mol Microbiol 5:1447-1457.

Broach JR, Hicks JB (1980) Replication and recombination functions associated with the yeast plasmid, 2 μ circle. Cell 21:501-508.

Caluwaerts S, Vandenbroucke K, Steidler L, Neirynck S, Vanhoenacker P, Corveleyn S, Watkins B, Sonis S, Coulie B, Rottiers P (2010) AG013, a mouth rinse formulation of *Lactococcus lactis* secreting human Trefoil Factor 1, provides a safe and efficacious therapeutic tool for treating oral mucositis. Oral Oncol 46:564-570.

Cello J, Paul AV, Wimmer E (2002) Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. Science 297:1016-1018.

Chang TL, Chang CH, Simpson DA, Xu Q, Martin PK, Lagenaur LA, Schoolnik GK, Ho DD, Hillier SL, Holodniy M, Lewicki JA, Lee PP (2003) Inhibition of HIV infectivity by a natural human isolate of *Lactobacillus jensenii* engineered to express functional two-domain CD4. Proc Natl Acad Sci USA 100:11672-11677.

Chen WY, Ho JW, Huang JD, Watt RM (2011) Functional characterization of an alkaline exonuclease and single strand annealing protein from the SXT genetic element of *Vibrio cholerae*. BMC Mol Biol 12:16.

Chen Y, Caruso L, McClane B, Fisher D, Gupta P (2007) Disruption of a toxin gene by introduction of a foreign gene into the chromosome of *Clostridium perfringens* using targetron-induced mutagenesis. Plasmid 58:182-189.

Cho SW, Kim S, Kim JM, Kim JS (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol 31:230-232.

Choi HJ, Ahn JH, Park S-H, Do KH, Kim J, Moon Y (2012) Enhanced wound healing by recombinant *Escherichia coli* Nissle 1917 via human epidermal growth factor receptor in human intestinal epithelial cells: therapeutic iplication using recombinant probiotics. Infect Immun 80:1079-1087.

Citorik RJ, Mimee M, Lu TK (2014) Sequence-specific antimicrobials using efficiently delivered RNA-guided

nucleases. Nat Biotechnol 32:1141-1145.
Cobb RE, Ning JC, Zhao H (2013) DNA assembly techniques for next-generation combinatorial biosynthesis of natural products. J Ind Microbiol Biotechnol 41:469-477.
Coleman JR, Papamichail D, Skiena S, Futcher B, Wimmer E, Mueller S (2008) Virus attenuation by genome-scale changes in codon pair bias. Science 320:1784-1787.
Costantino N, Court DL (2003) Enhanced levels of lambda Red-mediated recombinants in mismatch repair mutants. Proc Natl Acad Sci USA 100:15748-15753.
DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM (2013) Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res 41:4336-4343.
Dong H, Tao W, Gong F, Li Y, Zhang Y (2014) A functional recT gene for recombineering of *Clostridium*. J Biotechnol 173:65-67.
Duan F, March JC (2010) Engineered bacterial communication prevents *Vibrio cholerae* virulence in an infant mouse model. Proc Natl Acad Sci USA 107:11260-11264.
Ebina H, Misawa N, Kanemura Y, Koyanagi Y (2013) Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. Sci Rep 3:2510.
Ellis HM, Yu DG, DiTizio T, Court DL (2001) High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. Proc Natl Acad Sci USA 98:6742-6746.
Ellis T, Adie T, Baldwin GS (2011) DNA assembly for synthetic biology: from parts to pathways and beyond. Integr Biol (Camb) 3:109-118.
Enyeart PJ, Chirieleison SM, Dao MN, Perutka J, Quandt EM, Yao J, Whitt JT, Keatinge-Clay AT, Lambowitz AM, Ellington AD (2013) Generalized bacterial genome editing using mobile group II introns and Cre-lox. Mol Syst Biol 9:685.
Fabret C, Ehrlich SD, Noirot P (2002) A new mutation delivery system for genome-scale approaches in *Bacillus subtilis*. Mol Microbiol 46:25-36.
Fehér T (2014) Systematic Reduction of Microbial Genomes. In Singh V, and Govil JN (eds). Applied Synthetic Biology. Studium Press, Houston, TX, USA, 191-247.
Fehér T, Karcagi I, Blattner FR, Posfai G (2012a) Bacteriophage recombineering in the lytic state using the lambda red recombinases. Microb Biotechnol 5:466-476.
Fehér T, Bogos B, Méhi O, Fekete G, Csörgö B, Kovács K, Pósfai G, Papp B, Hurst LD, Pál C (2012b) Competition between transposable elements and mutator genes in bacteria. Mol Biol Evol 29:3153-3159.
Feng Z, Zhang B, Ding W, Liu X, Yang DL, Wei P, Cao F, Zhu S, Zhang F, Mao Y, Zhu JK (2013) Efficient genome editing in plants using a CRISPR/Cas system. Cell Res 23:1229-1232.
Foligne B, Dessein R, Marceau M, Poiret S, Chamaillard M, Pot B, Simonet M, Daniel C (2007) Prevention and treatment of colitis with *Lactococcus lactis* secreting the immunomodulatory *Yersinia* LcrV protein. Gastroenterology 133:862-874.
Frazier CL, San Filippo J, Lambowitz AM, Mills DA (2003) Genetic manipulation of *Lactococcus lactis* by using targeted group II introns: generation of stable insertions without selection. Appl Environ Microbiol 69:1121-1128.
Friedland AE, Tzur YB, Esvelt KM, Colaiacovo MP, Church GM, Calarco JA (2013) Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. Nat Methods 10:741-743.
Gibson DG, Smith HO, Hutchison CA, 3rd, Venter JC, Merryman C (2010a) Chemical synthesis of the mouse mitochondrial genome. Nat Methods 7:901-903.
Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6:343-345.
Gibson DG, Benders GA, Axelrod KC, Zaveri J, Algire MA, Moodie M, Montague MG, Venter JC, Smith HO, Hutchison CA, 3rd (2008a) One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. Proc Natl Acad Sci USA 105:20404-20409.
Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, Stockwell TB, Brownley A, Thomas DW, Algire MA, Merryman C, Young L, Noskov VN, Glass JI, Venter JC, Hutchison CA, 3rd, Smith HO (2008b) Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. Science 319:1215-1220.
Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, Benders GA, Montague MG, Ma L, Moodie MM, Merryman C, Vashee S, Krishnakumar R, Assad-Garcia N, Andrews-Pfannkoch C, Denisova EA, Young L, Qi ZQ, Segall-Shapiro TH, Calvey CH, Parmar PP, Hutchison CA, 3rd, Smith HO, Venter JC (2010b) Creation of a bacterial cell controlled by a chemically synthesized genome. Science 329:52-56.
Gomaa AA, Klumpe HE, Luo ML, Selle K, Barrangou R, Beisel CL (2014) Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. Mbio 5:e00928-13.
Gouzy A, Larrouy-Maumus G, Bottai D, Levillain F, Dumas A, Wallach JB, Caire-Brandli I, de Chastellier C, Wu TD, Poincloux R, Brosch R, Guerquin-Kern JL, Schnappinger D, Sorio de Carvalho LP, Poquet Y, Neyrolles O (2014) *Mycobacterium tuberculosis* exploits asparagine to assimilate nitrogen and resist acid stress during infection. PLoS Pathog 10:e1003928.
Gratz SJ, Cummings AM, Nguyen JN, Hamm DC, Donohue

LK, Harrison MM, Wildonger J, O'Connor-Giles KM (2013) Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. Genetics 194:1029-1035.

Grillot-Courvalin C, Goussard S, Huetz F, Ojcius DM, Courvalin P (1998) Functional gene transfer from intracellular bacteria to mammalian cells. Nat Biotechnol 16:862-866.

Grogan DW, Stengel KR (2008) Recombination of synthetic oligonucleotides with prokaryotic chromosomes: substrate requirements of the *Escherichia coli*/lambdaRed and *Sulfolobus acidocaldarius* recombination systems. Mol Microbiol 69:1255-1265.

Grundstrom T, Zenke WM, Wintzerith M, Matthes HW, Staub A, Chambon P (1985) Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis. Nucleic Acids Res 13:3305-3316.

Gutterson NI, Koshland DE, Jr. (1983) Replacement and amplification of bacterial genes with sequences altered in vitro. Proc Natl Acad Sci USA 80:4894-4898.

Hamilton CM, Aldea M, Washburn BK, Babitzke P, Kushner SR (1989) New method for generating deletions and gene replacements in *Escherichia coli*. J Bacteriol 171:4617-4622.

Heigwer F, Kerr G, Boutros M (2014) E-CRISP: fast CRISPR target site identification. Nat Methods 11:122-123.

Hoffman RM (2012) The preclinical discovery of bacterial therapy for the treatment of metastatic cancer with unique advantages. Expert Opin Drug Discov 7:73-83.

Horvath P, Barrangou R (2010) CRISPR/Cas, the immune system of bacteria and archaea. Science 327:167-170.

Horzempa J, O'Dee DM, Shanks RM, Nau GJ (2010) *Francisella tularensis* DeltapyrF mutants show that replication in nonmacrophages is sufficient for pathogenesis in vivo. Infect Immun 78:2607-2619.

Huang H, Zheng G, Jiang W, Hu H, Lu Y (2015) One-step high-efficiency CRISPR/Cas9-mediated genome editing in *Streptomyces*. Acta Bioch Bioph Sin 47:231-243.

Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JR, Joung JK (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol 31:227-229.

Isberg RR, Voorhis DL, Falkow S (1987) Identification of invasin: a protein that allows enteric bacteria to penetrate cultured mammalian cells. Cell 50:769-778.

Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol 31:233-239.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337:816-821.

Jurica MS, Stoddard BL (1999) Homing endonucleases: structure, function and evolution. Cell Mol Life Sci 55:1304-1326.

Kiro R, Shitrit D, Qimron U (2014) Efficient engineering of a bacteriophage genome using the type I-E CRISPR-Cas system. RNA Biol 11:42-44.

Kolisnychenko V, Plunkett G, 3rd, Herring CD, Feher T, Posfai J, Blattner FR, Posfai G (2002) Engineering a reduced *Escherichia coli* genome. Genome Res 12:640-647.

Kumar S, Smith KP, Floyd JL, Varela MF (2010) Cloning and molecular analysis of a mannitol operon of phosphoenolpyruvate-dependent phosphotransferase (PTS) type from *Vibrio cholerae* O395. Arch Microbiol 193:201-208.

Kusano K, Takahashi NK, Yoshikura H, Kobayashi I (1994) Involvement of RecE exonuclease and RecT annealing protein in DNA double-strand break repair by homologous recombination. Gene 138:17-25.

Lambowitz AM, Zimmerly S (2004) Mobile group II introns. Annu Rev Genet 38:1-35.

Lartigue C, Glass JI, Alperovich N, Pieper R, Parmar PP, Hutchison CA, 3rd, Smith HO, Venter JC (2007) Genome transplantation in bacteria: changing one species to another. Science 317:632-638.

Lartigue C, Vashee S, Algire MA, Chuang RY, Benders GA, Ma L, Noskov VN, Denisova EA, Gibson DG, Assad-Garcia N, Alperovich N, Thomas DW, Merryman C, Hutchison CA, 3rd, Smith HO, Venter JC, Glass JI (2009) Creating bacterial strains from genomes that have been cloned and engineered in yeast. Science 325:1693-1696.

Lee HY, Cho SA, Lee IS, Park JH, Seok SH, Baek MW, Kim DJ, Lee SH, Hur SJ, Ban SJ, Lee YK, Han YK, Cho YK (2007) Evaluation of phoP and rpoS mutants of *Salmonella enterica* serovar *Typhi* as attenuated typhoid vaccine candidates: virulence and protective immune responses in intranasally immunized mice. FEMS Immunol Med Microbiol 51:310-318.

Lin S-R, Yang H-C, Kuo Y-T, Liu C-J, Yang T-Y, Sung K-C, Lin Y-Y, Wang H-Y, Wang C-C, Shen Y-C, Wu F-Y, Kao J-H, Chen D-S, Chen P-J (2014) The CRISPR/Cas9 system facilitates clearance of the intrahepatic HBV templates in vivo. Mol Ther Nucleic Acids 3:e186.

Lu TK, Collins JJ (2007) Dispersing biofilms with engineered enzymatic bacteriophage. Proc Natl Acad Sci USA 104:11197-11202.

Lu TK, Collins JJ (2009) Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy. Proc Natl Acad Sci USA 106:4629-4634.

Luo ML, Mullis AS, Leenay RT, Beisel CL (2015) Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression. Nucleic Acids Res 43:674-681.

Marinelli LJ, Piuri M, Swigonova Z, Balachandran A, Oldfield LM, van Kessel JC, Hatfull GF (2008) BRED: a simple and powerful tool for constructing mutant and re-

combinant bacteriophage genomes. PLoS One 3:e3957.
Martel B, Moineau S (2014) CRISPR-Cas: an efficient tool for genome engineering of virulent bacteriophages. Nucleic Acids Res 42:9504-9513.
Miller VL, Mekalanos JJ (1988) A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires toxR. J Bacteriol 170:2575-2583.
Murphy KC (1998) Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. J Bacteriol 180:2063-2071.
Muyrers JP, Zhang Y, Testa G, Stewart AF (1999) Rapid modification of bacterial artificial chromosomes by ET-recombination. Nucleic Acids Res 27:1555-1557.
Nyerges A, Csoergo B, Nagy I, Latinovics D, Szamecz B, Posfai G, Pal C (2014) Conditional DNA repair mutants enable highly precise genome engineering. Nucleic Acids Res 42:e62.
Oh J-H, van Pijkeren J-P (2014) CRISPR-Cas9-assisted recombineering in *Lactobacillus reuteri*. Nucleic Acids Res 42:e131.
Oppenheim AB, Rattray AJ, Bubunenko M, Thomason LC, Court DL (2004) In vivo recombineering of bacteriophage lambda by PCR fragments and single-strand oligonucleotides. Virology 319:185-189.
Palonen E, Lindstrom M, Karttunen R, Somervuo P, Korkeala H (2011) Expression of signal transduction system encoding genes of *Yersinia pseudotuberculosis* IP32953 at 28 degrees C and 3 degrees C. PLoS One 6:e25063.
Pang Q-f, Ji Y, Bermudez-Humaran LG, Zhou Q-m, Hu G, Zeng Y (2009) Protective effects of a heme oxygenase-1 -secreting *Lactococcus lactis* on mucosal injury induced by hemorrhagic shock in rats. J Surg Res 153:39-45.
Pang Q, Ji Y, Li Y, Bermudez-Humaran LG, Hu G, Zeng Y (2008) Intragastric administration with recombinant *Lactococcus lactis* producing heme oxygenase-1 prevents lipopolysaccharide-induced endotoxemia in rats. FEMS Microbiol Lett 283:62-68.
Pennazio S (2006) The origin of phage virology. Riv Biol 99:103-129.
Perutka J, Wang W, Goerlitz D, Lambowitz AM (2004) Use of computer-designed group II introns to disrupt *Escherichia coli* DExH/D-box protein and DNA helicase genes. J Mol Biol 336:421-439.
Posfai G, Szybalski W (1988) Increasing the FokI cleavage specificity from 5 to 7 base pairs by two-step methylation. Nucleic Acids Res 16:6245.
Pósfai G, Kolisnychenko V, Bereczki Z, Blattner FR (1999) Markerless gene replacement in *Escherichia coli* stimulated by a double-strand break in the chromosome. Nucleic Acids Res 27:4409-4415.
Raman S, Rogers JK, Taylor ND, Church GM (2014) Evolution-guided optimization of biosynthetic pathways. Proc Natl Acad Sci USA 111:17803-17808.
Ranallo RT, Barnoy S, Thakkar S, Urick T, Venkatesan MM (2006) Developing live *Shigella* vaccines using lambda Red recombineering. FEMS Immunol Med Microbiol 47:462-469.
Rao S, Hu S, McHugh L, Lueders K, Henry K, Zhao Q, Fekete RA, Kar S, Adhya S, Hamer DH (2005) Toward a live microbial microbicide for HIV: commensal bacteria secreting an HIV fusion inhibitor peptide. Proc Natl Acad Sci USA 102:11993-11998.
Rodriguez SA, Davis G, Klose KE (2009) Targeted gene disruption in *Francisella tularensis* by group II introns. Methods 49:270-274.
Ruvkun GB, Ausubel FM (1981) A general method for site-directed mutagenesis in prokaryotes. Nature 289:85-88.
Sahu SN, Anriany Y, Grim CJ, Kim S, Chang Z, Joseph SW, Cinar HN (2013) Identification of virulence properties in *Salmonella Typhimurium* DT104 using *Caenorhabditis elegans*. PLoS One 8:e76673.
Sampson TR, Weiss DS (2014) CRISPR-Cas systems: new players in gene regulation and bacterial physiology. Front Cell Infect Microbiol 4:37.
Sander JD, Maeder ML, Reyon D, Voytas DF, Joung JK, Dobbs D (2010) ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. Nucleic Acids Res 38:W462-468.
Sawitzke JA, Costantino N, Li XT, Thomason LC, Bubunenko M, Court C, Court DL (2011) Probing cellular processes with oligo-mediated recombination and using the knowledge gained to optimize recombineering. J Mol Biol 407:45-59.
Schotte L, Steidler L, Vandekerckhove J, Remaut E (2000) Secretion of biologically active murine interleukin-10 by *Lactococcus lactis*. Enzyme Microb Technol 27:761-765.
Smith HO, Hutchison CA, 3rd, Pfannkoch C, Venter JC (2003) Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides. Proc Natl Acad Sci USA 100:15440-15445.
Steidler L, Hans W, Schotte L, Neirynck S, Obermeier F, Falk W, Fiers W, Remaut E (2000) Treatment of murine colitis by *Lactococcus lactis* secreting interleukin-10. Science 289:1352-1355.
Stemmer WP, Crameri A, Ha KD, Brennan TM, Heyneker HL (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. Gene 164:49-53.
Sternberg N, Hamilton D (1981) Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites. J Mol Biol 150:467-486.
Suda T, Liu D (2007) Hydrodynamic gene delivery: its principles and applications. Mol Ther 15:2063-2069.

Sun N, Abil Z, Zhao H (2012) Recent advances in targeted genome engineering in mammalian systems. Biotechnol J 7:1074-1087.

Thomason LC, Costantino N, Shaw DV, Court DL (2007) Multicopy plasmid modification with phage lambda Red recombineering. Plasmid 58:148-158.

Umenhoffer K, Fehér T, Baliko G, Ayaydin F, Pósfai J, Blattner FR, Pósfai G (2010) Reduced evolvability of *Escherichia coli* MDS42, an IS-less cellular chassis for molecular and synthetic biology applications. Microb Cell Fact 9:38.

van Kessel JC, Hatfull GF (2007) Recombineering in *Mycobacterium tuberculosis*. Nat Methods 4:147-152.

Vandenbroucke K, Hans W, Van Huysse J, Neirynck S, Demetter P, Remaut E, Rottiers P, Steidler L (2004) Active delivery of trefoil factors by genetically modified *Lactococcus lactis* prevents and heals acute colitis in mice. Gastroenterology 127:502-513.

Vandenbroucke K, de Haard H, Beirnaert E, Dreier T, Lauwereys M, Huyck L, Van Huysse J, Demetter P, Steidler L, Remaut E, Cuvelier C, Rottiers P (2010) Orally administered *L. lactis* secreting an anti-TNF Nanobody demonstrate efficacy in chronic colitis. Mucosal Immunol 3:49-56.

Wang HH, Xu G, Vonner AJ, Church G (2011) Modified bases enable high-efficiency oligonucleotide-mediated allelic replacement via mismatch repair evasion. Nucleic Acids Res 39:7336-7347.

Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM (2009) Programming cells by multiplex genome engineering and accelerated evolution. Nature 460:894-898.

Wild J, Hradecna Z, Pósfai G, Szybalski W (1996) A broad-host-range in vivo pop-out and amplification system for generating large quantities of 50- to 100-kb genomic fragments for direct DNA sequencing. Gene 179:181-188.

Woolston BM, Edgar S, Stephanopoulos G (2013) Metabolic engineering: past and future. Annu Rev Chem Biomol Eng 4:259-288.

Xiang S, Fruehauf J, Li CJ (2006) Short hairpin RNA-expressing bacteria elicit RNA interference in mammals. Nat Biotech 24:697-702.

Zhen S, Hua L, Takahashi Y, Narita S, Liu Y-H, Li Y (2014) In vitro and in vivo growth suppression of human papillomavirus 16-positive cervical cancer cells by CRISPR/Cas9. Biochem Biophys Res Commun 450:1422-1426.

Zhong J, Karberg M, Lambowitz AM (2003) Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. Nucleic Acids Res 31:1656-1664.