CHAPTER 3

Recombinant DNA Technology

SUMMARY

Isolation and purification of DNA are essential for many applications in biotechnology. The first steps of DNA isolation and purification involve removal of the cell wall, cell membrane, intracellular components, and DNA-associated proteins. Finally, the RNA must be degraded with RNA-specific enzymes called ribonucleases. The pure sample of DNA can finally be used in a variety of experiments.

Often in biotechnology applications, various fragment sizes exist in a single sample. Gel electrophoresis is a technique that separates DNA based on size. The simplest form of gel electrophoresis involves agarose, which is a substance derived from seaweed that is melted and solidifies when cooled. Wells are created in the molten agarose by adding a comb. When the agarose solidifies, the comb is removed, leaving behind a slight indentation that houses the initial DNA sample. DNA has an overall negative charge due to the phosphate functional groups. The DNA sample is added to the wells, and a current is applied across the agarose. DNA migrates through the agarose toward the positive electrode. The larger fragments of DNA do not move as quickly through the agarose as the shorter fragments, thus separating DNA on the basis of size. The DNA in the agarose is visualized under ultraviolet light using ethidium bromide, a fluorescent stain. Other gels can be used instead of agarose, depending on the size of the fragment to be separated. Polyacrylamide is useful for the separation of very small DNA fragments and also fragments that differ by only one nucleotide, such as in the case of DNA sequencing. For very large fragments, a technique called pulsed field gel electrophoresis (PFGE) is used. PFGE separates extremely large DNA fragments by alternating the current at right angles to increase the area of migration. Finally, in gradient gel electrophoresis, the concentration of acrylamide, buffer, or salts is altered across the gel to help eliminate secondary structures that can interfere with migration patterns in the gel.

DNA is cut into various fragment sizes by using enzymes called restriction endonucleases. These restriction enzymes recognize specific sequences on the DNA and cut both strands. There are two types of restriction enzymes. Type I restriction enzymes cut the DNA strands upwards of 1000 bases from the recognition sequence. Type II enzymes cut the DNA at the recognition sequence and are the most useful for biotechnology. Some restriction enzymes cut the DNA cleanly in two, producing a "blunt end" for each new fragment. Other enzymes make a staggered cut, which produces an overhang of a few bases on one strand from each new half. This is called a "sticky end." Two fragments of DNA that are cut with the same restriction enzyme can be ligated back together using DNA ligase, another enzyme that is very useful in biotechnology techniques such as gene cloning.

Many methods are used to detect DNA or RNA in a sample of an experiment. The total concentration of nucleic acid in a sample is often determined by measuring the absorbance of ultraviolet light. Specific DNA fragments are detected by labeling the nucleic acids with radioactive isotopes or fluorescent tags. The DNA is then detected by autoradiography or with photodetectors, respectively. Chemical tags also exist for similar application and are usually safer and less expensive to use than their radioactive or fluorescent counterparts. Detection for chemical tags relies on chemical reactions that produce insoluble blue precipitates or chemiluminescence.

Nucleic acid hybridization is a method to determine how closely related nucleic acid from two different sources is. Southern and Northern blots examine the relatedness of DNA and RNA, respectively, to a shorter piece of DNA, called a probe. The DNA probe has a known nucleotide sequence labeled with one of the tags discussed previously to enable detection. In a Southern blot, total DNA is isolated from a sample, digested with restriction enzymes, separated by size on a gel, and rendered single-stranded by addition of a strong acid. The single-stranded DNA within the gel is then transferred to a membrane, and the labeled DNA probe is added. The probe binds to complementary sequences on the membrane, which are then detected using a method appropriate to the tag. The specificity of the probe to the DNA on

the membrane can be altered by adjusting the temperature of hybridization. Northern blot is a similar process except that the target nucleic acid is usually mRNA that is also separated by size on a gel, although it is already single-stranded. Northern blots are more effective in testing eukaryotes because the presence of introns in DNA sequences makes hybridization more difficult in a Southern blot. Dot blots are variations to either Southern or Northern blots. In these tests, instead of separating the DNA by size, it is dropped directly onto the membrane as a dot, probed, and finally detected. Hybridization can also be performed *in situ*, meaning within a cell or thin section of tissue. Fluorescence *in situ* hybridization (FISH) involves labeling a probe with a fluorescent tag and hybridizing it to denatured DNA or RNA within a cell. Subtractive hybridization is useful to determine regions of genes that are missing, such as in the case of genetic defects. Also, subtractive hybridization can be used to identify genes that are expressed only under certain conditions.

Cloning vectors are small, circular pieces of DNA. Most often these DNA pieces are plasmids, but other DNA elements from viruses and even some that are artificially constructed are now being used. Cloning vectors are useful for many techniques, including sequencing of DNA, constructing probes for hybridization, cloning genes, and constructing gene libraries. Some useful traits of cloning vectors include small size, abundance within a cell, and ease in isolating. Also, they must be easy to manipulate and detect. Genes for antibiotic resistance are often useful in selecting for and detecting the presence of a cloning vector within a bacterial culture. A region on the vector called a polylinker contains numerous restriction enzymes sites that are used to insert foreign pieces of DNA.

Bacterial cloning vectors typically hold smaller pieces of target DNA. Some vectors can also be derived from bacteriophage, which are viruses that infect bacteria only. Cosmids are viral genomes that have some genes removed in order to package more target DNA into them. Artificial chromosomes hold yet the largest amount of DNA and are specific to the host: yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), and P1 bacteriophage artificial chromosomes (PACs).

Once the cloning vectors are constructed, they are inserted into host cells. The process of inserting foreign, "naked" DNA into a cell is called transformation. Transformation requires that the host cell be made competent to uptake the DNA, either by generating holes in their cell walls and membranes or by subjecting them to a high-voltage shock. Some vectors are incompatible, meaning that they cannot be within the same cell due to constraints regarding replication. Often, the cell will lose one of two vectors if they are deemed incompatible.

Whole genomes of organisms can be digested by restriction enzymes and inserted into cloning vectors to construct gene and expression libraries. Both types of libraries have many uses, including sequencing, identifying new genes, comparing the genes of different organisms for gene libraries, and examining protein products from genes in expression libraries. The vectors used in expression libraries have a few more features than cloning vectors. Most important is the promoter region. The protein product of a foreign gene is often toxic to *E. coli*, the host of choice for most biotechnology techniques. Controlling the expression of the foreign gene is crucial and relies on the presence of a promoter than can be induced or de-repressed when necessary. Additionally, expression vectors have a small sequence that encodes a tag. This tag, when translated, produces a short piece of polypeptide that is attached to the protein product of the target gene. This allows a researcher to identify and isolate the specific protein product from the cloned gene.

Recombineering is a technique that utilizes homologous recombination to create large recombinant DNA vectors. Engineered bacterial host strains express lambda phage RED enzymes, which recognize homologous sequences and recombine them into a single DNA molecule. This process is utilized when genes are long or restriction enzyme choices are limited.

Gateway® cloning is a relatively newer cloning system that uses lambda phage integration and excision sites instead of polylinkers. Lambda enzyme integrase creates staggered cuts at

the centers of attachments sites in the bacterial host and in the phage, called *attB* and *attP*, respectively. The two ends can be recombined, which incidentally creates two new sites with different sequences than the originals called *attL* and *attR*. Excisionase is a second lambda phage enzyme that recognizes these new sites and removes the integrated DNA. The Gateway® cloning vectors use the lambda integration and excision system. DNA of interest is first cloned into an entry vector, which creates *attL* sites on either end of the DNA segment. After this initial cloning, two reactions called the LR reaction and BP reaction move the gene of interest to the destination vector. Bacteria carrying the clone survive because the selection process involves replacement of the gene of interest in the entry clone with a gene that is expressed into a toxin. Only those cells carrying the gene of interest survive.

Case Study Marker-Free Plasmids for Biotechnological Applications—Implications and Perspectives

Pedro H. Oliveira, Juergen Mairhofer (2013). *Trends in Biotechnology* 13(9), 539–547.

During traditional gene cloning and transformation, antibiotic resistance is often used as a selective marker for host cells harboring the plasmid. The use of antibiotics as a selecting agent for plasmid DNA constitutes a potential safety risk if the resulting plasmid DNA is designated for DNA vaccines. Additionally, antibiotic resistance markers contribute to structural plasmid instabilities and decreased gene delivery efficiency.

The authors of this review provide insight into the various marker-free plasmids that have been developed and discuss the implications in not only medicine but also other biotechnological applications.

Transduction of viruses containing corrected versions of genes into human host cells could potentially be used as gene therapy for genetic disorders. However, nonviral gene therapies, specifically using plasmid-based delivery systems, have gained popularity. What are the advantages and disadvantages of using a plasmid-based delivery system over viral therapies?

Since plasmid DNA vectors are not infectious agents like viruses are, one advantage of using plasmid DNA is safety. Other advantages include stability, lower toxicity, and ease of production. The expression of the transgene within nonviral systems is typically weaker. This disadvantage of plasmid DNA has encouraged the improvement of plasmid-based systems, including their delivery.

How have plasmid-based systems been improved?

The uptake, transport, and expression of plasmid-based transgenes have been improved through modifications to specific elements contained within the vector. For example, to decrease the susceptibility of the DNA to nucleases and increase expression of the transgene, specific polyadenylation (polA) sequences were modified. Additionally, extraneous sequences were eliminated, and codon usage in the host was optimized. Traditional plasmid DNA methodologies use antibiotic resistance markers as a selecting agent for the plasmid. Antibiotic resistance-free selection systems were also developed as part of the system improvements.

What types of antibiotic resistance-free selection systems have been developed?

Current improvements to plasmid-based systems have created antibiotic-free selective methods. These systems include the use of minicircles, bacterial backbone-reduced plasmids, operator-repressor titration, RNA/RNA interaction, and mini-intronic plasmids. Minicircles are supercoiled DNA cassettes that do not contain any bacterial plasmid DNA, including the antibiotic cassette. They are created by recombination of target sites surrounding the transgene either using recombinase, integrase, or resolvase. The construction is laborious and currently limited to small-scale production. Bacterial backbone-reduced plasmids include pCOR/pFAR systems. Both of these systems rely on the suppression of an amber nonsense mutation. This mutation is within an essential chromosomal gene such as *argE* or *thyA*. The plasmid contains the amber nonsense suppressor tRNA. If the plasmid is not present, the suppressor tRNA is not expressed, and the mutation in the essential gene is not suppressed. Any host cell not harboring the plasmid does not grow. An operator-repressor titration system requires a bacterial essential gene, such as for cell wall biosynthesis, that is regulated by an inducible promoter-operator. The repressor is titrated from the gene's operator upon transformation of a plasmid containing the operator site. HIV-1 DNA vaccine studies have utilized this approach. During RNA/RNA interaction, antisense RNA regulators prevent antibiotic resistance genes from being used. Some of the RNA/RNA interaction technologies are currently in clinical trials, and others are being evaluated *in vitro* human mesenchymal and mouse neural stem cells.

Why is biosafety a major concern with using antibiotic resistance markers in biotechnology? In addition to biosafety, what are other concerns associated with antibiotic resistance markers?

The use of antibiotics in all aspects of biotechnology, including with genetically engineered plants, allows the possibility of accidental release of these resistance genes into the environment. Antibiotic resistance genes could by spread by horizontal gene transfer. Pathogens that acquire these resistance genes could become multidrug resistant, posing much risk to human health. Antibiotic resistance markers also encompass a fairly decent portion of plasmid DNA. Eliminating these genetic markers from the DNA increases stability of the plasmid and efficiency of transformation. Additionally, expression of antibiotic resistance genes can alter the gene expression profiles of host cells, including human cell lines, to a point that their use is discouraged for the production of stable cell lines. Vertebrate immune systems even recognize portions of the bacterial-derived plasmid DNA backbone and elicit an inflammatory response that drives an adaptive immune response that is appropriate only if the intent is a DNA vaccine. However, if expression is desired, an antibiotic marker-free plasmid or CpG-depleted vector should be used instead. Lastly, bacterially derived backbone sequences are transcriptionally silenced when introduced into mammalian cells. Evidence suggests the reason this occurs is that some sequences of the vector resemble mammalian heterochromatin and the vector is large enough for histone modifications to occur.

(*Continued*)

Case Study Marker-Free Plasmids for Biotechnological Applications—Implications and Perspectives—cont'd

How would using antibiotic marker-free vectors affect the various biotechnology applications?

Much controversy stems from the development and use of transgenic agriculture plants. Many consumers are fearful of accidental release of the antibiotic resistance genes into the environment or to other organisms. The regulatory process when using these selective markers is also lengthy. Using antibiotic-free selectable markers would greatly simplify the regulatory process and improve consumer relations. Additionally, using nonintegrating plasmid DNA as a possible means to treat genetic disorders in human stem cells is safer than using viral vectors that could potentially integrate and disrupt tumor suppressor genes, activate oncogenes, or silence portions of the genome.

The use of antibiotic selection systems in recombinant DNA technology and biotechnology poses many challenges. Several drawbacks exist for the use of antibiotics in this field, including biosafety concerns, public relations, regulatory issues, as well as delivery, expression, and stability of the vectors containing these antibiotics markers in the host cells. With these challenges in mind, researchers have developed several different types of selections systems that do not use antibiotic resistance. In this review, these authors have discussed the use of antibiotics, the motivations for developing antibiotic-free marker systems, and also several of these new systems.

Review

Marker-free plasmids for biotechnological applications – implications and perspectives

Pedro H. Oliveira[1,2] and Juergen Mairhofer[3]

[1] Institut Pasteur, Microbial Evolutionary Genomics, Département Génomes et Génétique, F-75015 Paris, France
[2] CNRS, UMR3525, F-75015 Paris, France
[3] Institute of Applied Microbiology, University of Natural Resources and Life Sciences, Vienna, Austria

Nonviral gene therapy and DNA vaccines have become the first promising approaches to treat, cure, or ultimately prevent disease by providing genetic information encoded on a plasmid. Since 1989, more than 1800 clinical trials have been approved worldwide, and approximately 20% of them are using plasmid DNA (pDNA) as a vector system. Although much safer than viral approaches, DNA vectors generally do encode antibiotic resistance genes in the plasmid backbone. These antibiotic resistance markers constitute a possible safety risk, and they are associated with structural plasmid instabilities and decreased gene delivery efficiency. These drawbacks have initiated the development of various antibiotic marker-free selection approaches. We provide an overview on the potential implications of marker-free plasmids and perspectives for their successful biotechnological use in the future.

Plasmid vector design for DNA vaccine or gene therapy applications

Plasmid-based gene delivery (see Glossary) has gained widespread and mounting attention in the last two decades, in particular due to its promising potential in the correction of genetic defects and in the prevention or treatment of infectious diseases and cancer. When confronted with their viral counterparts, plasmid DNA (pDNA) vectors offer appealing advantages, particularly in terms of safety, stability, lower toxicity, and easier scalability of production. However, transgene expression attained with nonviral alternatives is typically weaker and poorly sustained. Therefore, it is of the utmost importance to improve current pDNA vector systems and their mode of delivery.

A therapeutic pDNA molecule is typically built around a modular structure, comprising a unit essential for propagation in a bacterial host (usually *Escherichia coli*) and a eukaryotic transcription unit (Figure 1). Remarkable improvements in the efficacy of these molecules have in many cases arisen by tinkering with different elements of its structure. These include, for example, the modification

Corresponding author: Mairhofer, J. (juergen.mairhofer@boku.ac.at).
Keywords: plasmid vector design; antibiotic resistance; DNA vaccines; gene therapy; genetically modified plants; stem cells..

0167-7799/$ – see front matter
© 2013 Elsevier Ltd. All rights reserved. http://dx.doi.org/10.1016/j.tibtech.2013.06.001

Glossary

Amber suppressor tRNA: tRNA whose anticodon has been changed in order to recognize the unique nonsense codon 5′-UAG-3′ (amber) in the mRNA and insert an alternative amino acid at that position in the polypeptide chain.
Cytosine-phosphate-guanine (CpG) dinucleotides: immunostimulatory motifs found at high frequency and in an unmethylated state in bacterial DNA, whereas in mammalian DNA they are underrepresented and methylated at the C-5 position of cytosines. The base context in which the CpG motifs are found, as well as their number and spacing, are factors known to differentially stimulate the immune response.
Horizontal gene transfer (HGT): also known as lateral gene transfer, comprises the movement of genetic material between more or less related organisms by means other than those taking place during cell division (vertical gene transfer). In bacteria, HGT may occur via transformation, transduction, or conjugation mechanisms, and has been recognized as a significant cause of increased antibiotic resistance.
Minicircles: small, supercoiled minimal circular plasmids devoid of any prokaryotic backbone, including origin of replication and antibiotic resistance gene. They are generated *in vivo* by site-specific recombination.
Mini-intronic plasmids (MIPs): a novel class of pDNA vectors where the essential bacterial elements for plasmid replication and selection are placed within an engineered intron in the eukaryotic expression cassette. MIPs are able to overcome the transgene silencing observed with conventional plasmids.
***murselect* technology**: plasmid maintenance system that relies solely on the presence of a ColE1-like origin of replication. Plasmid selection and maintenance is achieved by functionally linking the RNAI encoded on the origin of replication to an artificial repressor protein encoded on the host genome that further controls the expression of an essential gene, whose promoter has been engineered to contain an operator binding site for the artificial repressor.
Operator–repressor titration (ORT): plasmid maintenance system that relies on the titration of repressor molecules of an essential chromosomal gene by an operator sequence located in a multicopy plasmid.
pCOR: this class of small (2 kb) plasmid vectors contains a minimal amount of bacterial DNA and an R6K-derived conditional origin of replication (COR). They are propagated in a specifically engineered *Escherichia coli* host strain carrying an amber mutation in the acetylornithine deacetylase gene, which is corrected by a suppressor tRNA encoded on the plasmid. This eliminates the need for antibiotic resistance genes and greatly reduces the possibility of plasmid dissemination into the environment.
pFAR: plasmids free of antibiotic resistance (pFAR) markers, which, like pCOR, rely on the suppression of a chromosomal amber mutation. In this case, the pFAR carries an amber suppressor tRNA that restores the normal growth of a thymidine auxotrophic strain.
Plasmid DNA (pDNA) vectors: small (usually between 3 and 7 kb), self-replicating circular double-stranded DNA molecules, whose backbone has been engineered in order to be used as vehicles for cloning and transfer of DNA fragments into a host cell. They typically contain a minimum set of features comprising a replication origin compatible with the host, a gene of interest (GOI) to be expressed, a promoter element capable of efficient and robust transcriptional activity, and a selection marker.
Plasmid-based gene delivery: introduction of therapeutic genes into designated target cells using pDNA as carrier.
RNA-OUT/RNA-IN: selection marker based on two antisense RNA regulators, RNA-IN and RNA-OUT, which are generally involved in the regulation of the IS*10* transposase activity. When used for the purpose of antibiotic resistance-free plasmid selection, the 150 bp RNA-OUT is cloned in the plasmid, while a lethal gene is placed in the chromosome under the control of the RNA-IN promoter.

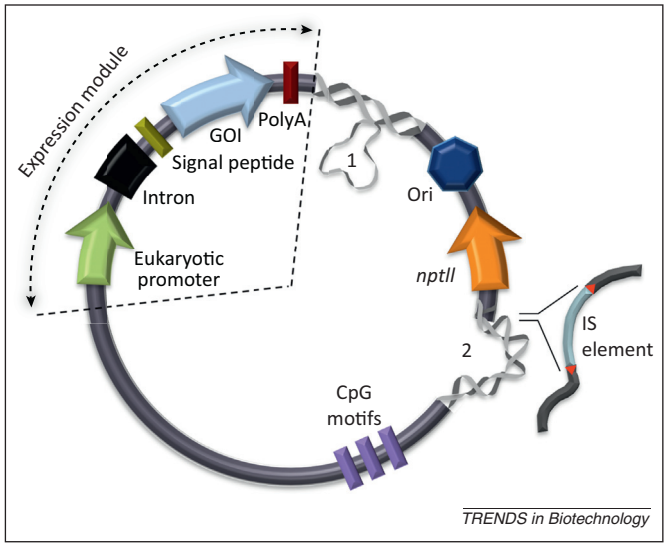

Figure 1. Structural elements and possible detrimental sequence motifs on a conventional therapeutic plasmid. Structural elements can be organized into (i) a bacterial propagation unit and (ii) a eukaryotic expression unit. The bacterial expression unit typically contains an origin of replication for the amplification of the plasmid in *Escherichia coli* (normally, theta-replicating, non-protein coding origins such as ColE1 are preferred over others) and an antibiotic resistance gene (often *nptII*) used for the selection of plasmid-bearing cells. The average bacterial propagation unit has a size of approximately 1.8 kb. The eukaryotic expression unit contains a promoter sequence, preferably tissue-specific, a 5′ untranslated region (5′UTR) including introns, the therapeutic gene of interest (GOI), and a polyadenylation signal (polyA). Also shown are examples of elements prone to undergo structural rearrangements. The situation marked as **1** illustrates the adoption of a slipped-misaligned structure typical of the presence of direct repeats. Such noncanonical (non-B) elements can ultimately halt fork progression and lead to deletion-formation events. The situation marked as **2** depicts the spontaneous transposition of a mobile element (insertion sequence, IS) into an intrinsically bent region upstream of the *nptII* gene. This occasionally generates a hybrid promoter between a -35 box carried by the IS element itself and a resident -10 box. Spontaneous insertions of IS elements may nevertheless occur at additional positions, in a process that is highly dependent on topological constraints of the target region [79] and abrupt DNA compositional asymmetries (P. Oliveira, unpublished data).

of polyadenylation (polyA) sequences to further increase transgene expression and resistance to nucleases [1], alterations in the type and load of cytosine-phosphate-guanine (CpG) motifs [2], changes in codon bias [3], and the elimination of extraneous sequences [4]. These differences can have profound effects on the cellular uptake of pDNA molecules, intracellular stability, nuclear transport, and consequently duration of transgene expression.

Current selection methods and alternative antibiotic resistance-free selection systems

In addition to the sequence necessary for replication, conventional plasmids contain a selectable marker that ensures selection during the cloning process and stable inheritance of pDNA during bacterial growth. For this purpose, pDNAs carry genes that confer resistance to a drug, mostly antibiotics. Currently, the neomycin/kanamycin resistance gene (neomycin phosphotransferase II, *nptII*), initially isolated from the transposon Tn5, is the most widely used selection marker in bacteria, plants, and mammalian cells because it is the only one tolerated by regulatory authorities [5]. The gene codes for the enzyme aminoglycoside-3′-phosphotransferase, which inactivates a range of aminoglycoside antibiotics by phosphorylation (e.g., kanamycin and G418). However, concerns about the spread of antibiotic resistance genes by horizontal gene transfer (HGT; e.g., to the microbiome of the vaccinated host) or the possibility of integration into the host's genome have spurred development of alternative selection mechanisms.

Minicircles

Minicircles are double-stranded, supercoiled expression cassettes devoid of the bacterial pDNA backbone. They are generated in *E. coli* by recombination between two recombinase target sites, flanking the eukaryotic expression unit. The recombination event generates the desired minicircle together with the so-called miniplasmid, encoding the undesired antibiotic resistance gene and the bacterial origin of replication (Figure 2A). Several different strategies for the generation of minicircles exist using either λ integrase [6], Cre recombinase from bacteriophage P1 [7], integrase from the *Streptomyces* bacteriophage ϕC31 [8] or Par resolvase from the multimer resolution system of plasmid pRK2 [9]. The manufacturing of minicircles is a laborious process that currently lacks reproducibility at the large scale. The most recent developments in the field of minicircles are reviewed elsewhere [10].

Bacterial backbone-reduced plasmids

pCOR/pFAR systems. The pCOR (plasmids with conditional origin of replication) and the pFAR (plasmids free of antibiotic resistance) selection systems rely on the suppression of an amber nonsense mutation introduced into an essential chromosomal bacterial gene. For that purpose, the plasmids contain an amber suppressor tRNA (~100 bp in size) that inserts a certain amino acid instead of terminating translation. The host strain for propagation of pCOR or pFAR has been modified to carry an amber stop codon in an essential gene that can only be complemented by the plasmid carrying the corresponding suppressor tRNA (Figure 2B). pCOR was designed to overcome an arginine auxotrophy, caused by introduction of an amber mutation into the acetylornithine deacetylase gene (*argE*), by replacing the antibiotic resistance gene for a synthetic amber suppressor tRNA [11]. In addition, the pCOR plasmid family utilizes an R6K-derived conditional origin of replication that requires an *E. coli* host strain that expresses a trans-acting π initiator protein encoded by a gene (*pir*) stably integrated into the chromosome. This containment strategy significantly reduces the risk of plasmid dissemination to other bacteria. pCOR plasmids have been produced with reasonable yields [12] and several preclinical and clinical studies have been already performed [13–15].

In contrast to pCOR, plasmids of the pFAR series harbor a high copy number pUC origin of replication. Here the amber mutation was introduced into the *thyA* gene, necessary for DNA synthesis and encoding for thymidylate synthase, consequently leading to a thymidine auxotrophic strain. Again, the complementation suppression is achieved by introduction of a suppressor tRNA on the plasmid backbone. *In vivo* and *in vitro* data for a pFAR plasmid encoding the luciferase reporter gene is available for mouse muscle, skin, and tumor cells [16].

Operator–repressor titration. The pORT technology relies on the phenomenon of repressor titration. Operator–repressor titration (ORT) requires a genetically modified

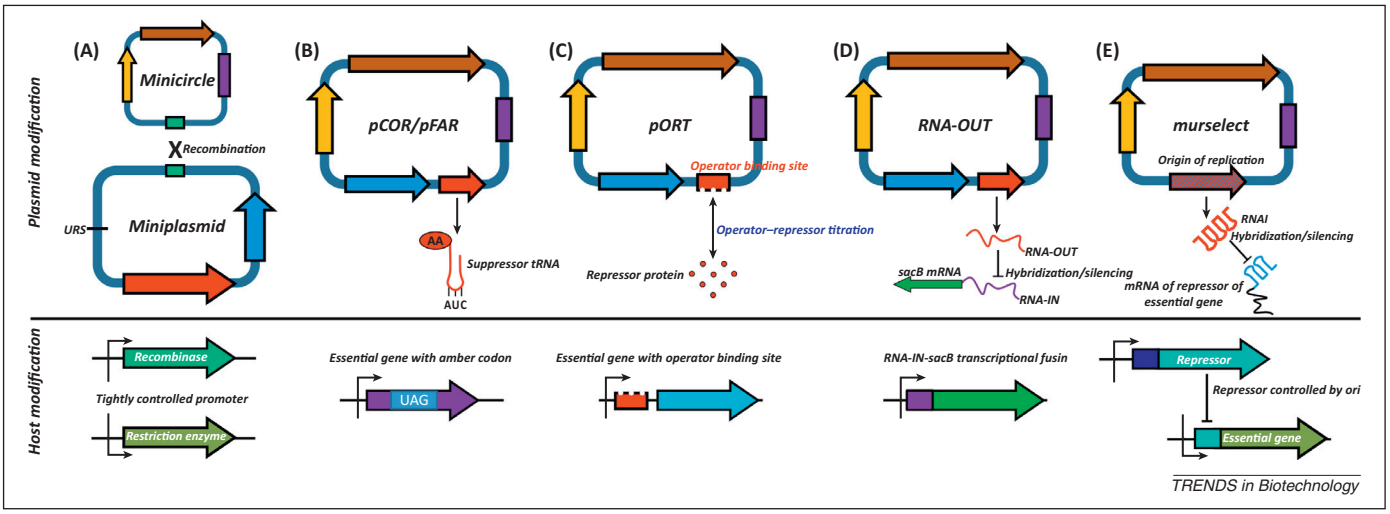

Figure 2. Overview of available antibiotic-resistance-free plasmid selection systems showing necessary plasmid and *Escherichia coli* host modifications. **(A)** Minicircle technology: minicircle and miniplasmid are resolved from a parental plasmid by recombination. The miniplasmid can be removed by cleavage with a unique restriction enzyme (URS). A suitable *E. coli* host for minicircle production must provide tightly controlled transcription of the recombinase gene and the URS gene. **(B)** Plasmid with conditional origin of replication (pCOR)/plasmid free of antibiotic resistance (pFAR): these plasmids encode a suppressor tRNA that can overcome an amber codon (TAG) that was introduced in an essential gene on the *E. coli* chromosome. **(C)** Plasmid with operator–repressor titration (pORT): multiple operator binding sites are integrated into the multicopy plasmid backbone, leading to a large excess of the plasmid-encoded operator compared to the chromosomal-encoded operator binding site that controls an essential gene. Titration of the repressor by the plasmid leads to expression of the essential gene. **(D)** RNA-OUT: a counter-selectable marker (*sacB*) is chromosomally integrated and constitutively expressed under the control of the RNA-IN promoter/leader. The *sacB* gene encodes a levansucrase, which is toxic for the cell in the presence of sucrose. When a 150 bp RNA-OUT antisense RNA is expressed from the plasmid, it will bind RNA-IN and repress the expression of *sacB*, allowing growth in the presence of sucrose. **(E)** *murselect* system: the plasmid only has to encode a ColE1-type origin of replication. Selection and maintenance is achieved by functionally linking the RNAI encoded on the origin of replication of the plasmid, to a repressor protein encoded on the host genome that further controls the expression of an essential gene. Key: Red arrow, element conferring selection; blue arrow, origin of replication; red/blue hatched arrow, element combining origin of replication and selection; yellow arrow, eukaryotic promoter; brown arrow, therapeutic gene; purple box, polyadenylation signal; green box, recombinase attachment site.

(GM) *E. coli* strain wherein an essential gene is regulated by an inducible promoter–operator [17] (Figure 2C). When this strain is transformed with a multicopy plasmid containing an operator sequence, the repressor is titrated, allowing expression of the essential tetrahydrodipicolinate *N*-succinyltransferase gene (*dapD*, involved in cell wall synthesis), and consequently cell growth [18]. pORT backbones have been used in some preclinical and clinical HIV-1 DNA vaccine studies [19–21], as well as for plasmid stabilization in *Salmonella enterica* for the generation of live bacterial vaccines [22–24].

Selection based on RNA/RNA interaction. Other plasmid selection systems use RNA antisense regulators to prevent the usage of antibiotic resistance genes. These approaches rely on a genetically engineered *E. coli* host strain and an RNA antisense regulator encoded by the plasmid that secures cell growth. Plasmids encoding the 150 bp RNA-OUT antisense RNA are capable of repressing the expression of a counter-selectable marker *sacB* (encoding for levansucrase), encoded on the host chromosome. Expression of *sacB* is toxic in the presence of sucrose and therefore plasmid selection is achieved in the presence of the aforementioned [25] (Figure 2D). Another approach exploits RNAI, a regulatory RNA naturally encoded on the origin of replication of ColE1-like plasmids (e.g., pUC19), to inhibit expression of a repressor protein that further controls transcription of the essential gene *murA* (encoding for uridine diphosphate *N*-acetylglucosamine enolpyruvoyl transferase, catalyzing the first committed step in the assembly of peptidoglycan) in the GM *murselect E. coli* strains [4]. This approach allows the selection of plasmids without the addition of other extrinsic elements to the plasmid backbone (Figure 2E). Both systems are scalable and allowed the production of pDNA in gram-scale quantities in fed batch fermentations [25,26]. *In vivo* data for RNA-OUT plasmids expressing influenza H5 hemagglutinin is available [25]. In addition, RNA-OUT vectors have already successfully completed nonclinical toxicology evaluation and some applications have recently transitioned from preclinical to clinical status. Marker-free plasmids, selected using the *murselect* technology, have been recently evaluated *in vitro* using human mesenchymal and mouse neural stem cells [27] (see *Perspectives* section).

New development: MIPs. In a recent publication, the development of a mini-intronic plasmid (MIP) vector system was reported [28]. Here, the essential bacterial backbone elements are placed within an engineered intron contained within the eukaryotic expression cassette. These specific plasmids utilize the antibiotic-free RNA-OUT selectable marker vector systems in combination with a high copy number pUC origin. The authors also generated MIPs with the commonly used kanamycin resistance gene, instead of the RNA-OUT technology, and observed that transgene expression levels were two magnitudes lower compared to conventional plasmids. For the MIP plasmid, it is particularly important to only contain a very short or no marker gene at all because these sequences provide potential alternative splicing sites (Figure 3). Overall, MIP vectors showed a two- to tenfold increase in transgene expression levels when compared to minicircles after transfection into quiescent tissue *in vivo* and *in vitro*. To date, the detailed mechanism of MIP-generated enhanced expression remains elusive, but the authors do propose an intron-mediated transcriptional enhancement due to

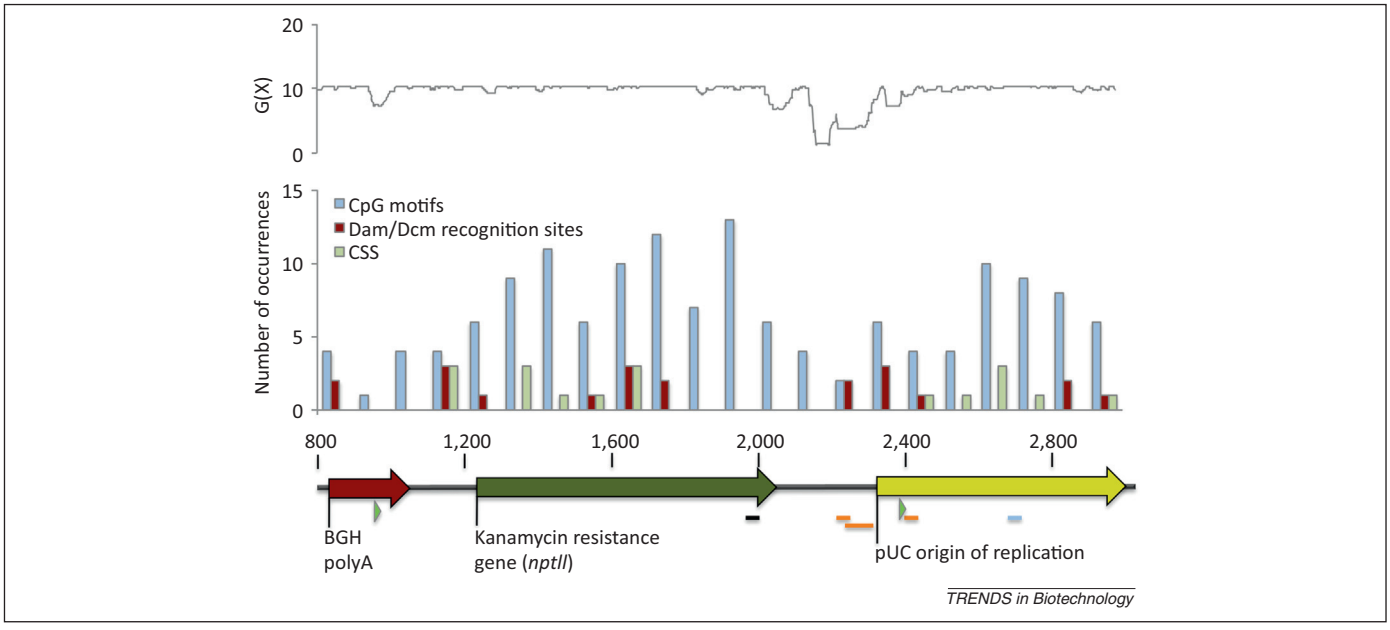

Figure 3. Abundance of potentially detrimental motifs present in the pVAX1 (Invitrogen, Carlsbad, CA, USA) region comprising the bovine growth hormone (BGH) polyadenylation (polyA), the *nptII* gene (Tn5 derived), and the origin of replication. The upper graph refers to the stress-induced duplex destabilization energy [G(x), kcal/mol], i.e., the energy cost that leads to strand separation in conditions of negative DNA superhelicity (see http://benham.genomecenter.ucdavis.edu/sibz/). Values of G(x) close to 10.2 correspond to full stability. The results show the presence of a slightly destabilized region within the BGH polyA, and a more prominent one encompassing the region upstream of the origin of replication and its 5′ extremity. Such regions are good candidates for nuclease attack due to their increased propensity to single strandness, as has been shown experimentally [80] (green arrows in the bottom scheme indicate the precise locations of these nuclease-prone regions). The middle plot shows the distribution of CG dinucleotides (CpG motifs, blue bars), Dam/Dcm recognition sites (red bars), and cryptic splice sites (CSS, green bars). The latter were computed using two different tools (NNSPLICE, http://www.fruitfly.org/seq_tools/splice.html; NETGENE2, http://www.cbs.dtu.dk/services/NetGene2/ with default parameters), and only the common hits were accepted as valid. The numeration shown is in accordance with the information provided by the manufacturer. It is interesting to note the particularly high abundance of these types of elements within the kanamycin gene and the origin of replication. For example, only the kanamycin resistance gene harbors a total of 74 CpG motifs, which if methylated, may exert a spreading effect into neighbor elements, adversely affecting their expression in *cis*. Dam and Dcm methylase recognition sites were also found, albeit at a lower proportion. Dam methylates position 6 of A in GATC (6meA), which is known to be prone to depurinations and transversions. By contrast, Dcm methylates the internal cytosine residue of the sequence CCWGG at position 5 (5meC), which is known to occasionally mutate to thymine. We have also looked at the presence of noncanonical (non-B) elements using the non-B database tool (http://nonb.abcc.ncifcrf.gov/apps/site/default). Only perfect repeats with sizes $\geq$10 bp were considered as potentially detrimental from the point of view of structural instability. A 10-bp long direct repeat pair was detected close to the 3′ extremity of the *nptII* gene (black line), whereas three 10- and 11-bp long inverted repeats (orange lines) were detected in the 5′ region of the origin of replication. The latter are possibly involved in the abovementioned duplex destabilization events. Moreover, one of these inverted repeat pairs closely matches the nuclease-prone region (green arrow). Finally, one motif prone to form Z-DNA (blue line) has been found within the origin of replication.

promoter proximal splice sites [29], which compensates for the slightly lower transformation efficiencies when compared to smaller minicircles (see Figure I in Box 1). Altogether, MIP vectors present an attractive alternative to minicircles owing to their superior expression profiles and the relatively simple production process, identical to the production of conventional pDNA. A schematic overview on the configuration of functional elements on a MIP vector can be found in Figure 4.

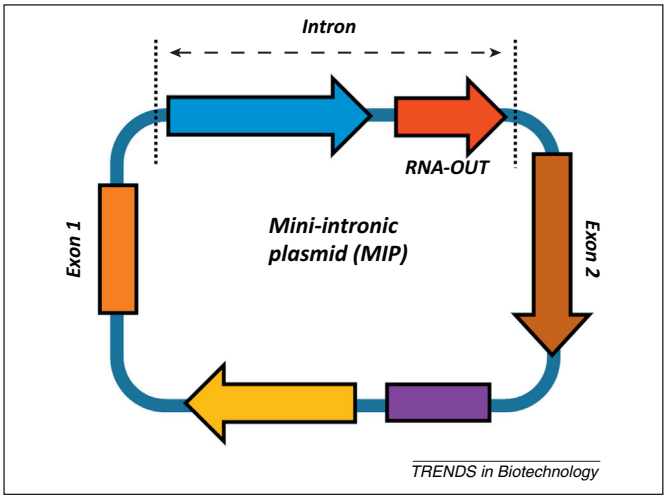

Figure 4. Schematic overview of the configuration of vector elements on a mini-intronic plasmid (MIP). The bacterial backbone elements (blue and red arrows) are placed within an intron sequence that is spliced during RNA processing. Key: Red arrow, element conferring selection; blue arrow, origin of replication; yellow arrow, eukaryotic promoter; orange box/exon 1, untranslated exon; brown arrow/exon 2, therapeutic gene; purple box arrow, polyadenylation signal.

Advantages of antibiotic resistance-free plasmids and minicircles

Biosafety

The use of antibiotic-resistance genes in plasmids for biotechnological applications causes potential biosafety concerns, such as the possible spread of resistance genes by HGT or the emergence of multidrug-resistant pathogens. The major causes for the increasing threat of multidrug-resistant bacteria are definitely the overuse of antimicrobial agents in hospitals and farms (according to the Pew charitable trust, 80% of the US antibiotic supply was used for livestock in 2011; www.saveantibiotics.org), as well as poorly controlled infection control practices [30]. However, according to a recent study, resistance genes sourced from synthetic plasmids released into the environment may represent a yet unrecognized source of antibiotic resistance [31]. Therefore, the use of antibiotic marker-free plasmids will eliminate possible biosafety concerns.

Currently, the *nptII* gene is the most commonly used selection marker because it is the only one tolerated by

Box 1. Plasmid size detrimentally affects gene delivery efficiency

In order to achieve transgene expression, a plasmid has to be routed through the crowded environment of the cytoplasm, to finally reach the nucleus. Diffusion of DNA in the cytoplasm of cells is strongly size dependent, with little or no diffusion for DNAs > 2000 bp [81]. It is believed that binding effects to cytoskeletal elements (which can act as molecular sieves) are not primarily responsible for the slowed diffusion of DNA molecules. The major problem seems to be molecular crowding and collisional interactions exerting a strong dependence on intracellular diffusion. The transfection of nondividing smooth muscle cells by a minicircle 2.9 kb in size is 77 times more efficient than a plasmid 52.5 kb in length [82]. A systematic study on the effect of plasmid size on promoter/enhancer activity in transient transfection assays revealed that expression declined as a function of vector size and that the sharpest decrease in reporter gene activity occurred when a 5.1 kb plasmid was increased by 0.65 kb [83]. Notably, this critical threshold appears at a very common plasmid size of approximately 6 kb. Overall, it seems that reduction in pDNA size enhances the transformation efficiencies and is therefore beneficial for pDNA delivery (see also Figure I).

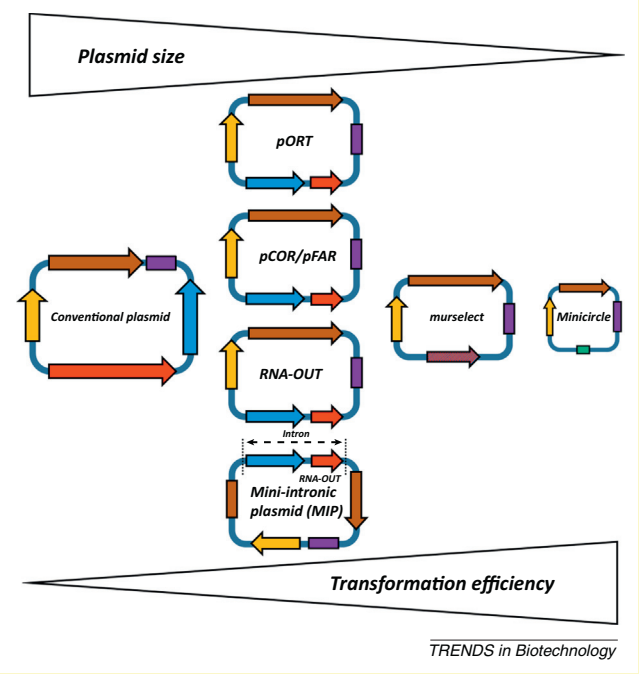

Figure I. Ranking of alternative plasmid selection approaches according to plasmid size and transformation efficiency. Key: Red arrow, element conferring selection; blue arrow, origin of replication; red/blue hatched arrow, element combining origin of replication and selection; yellow arrow, eukaryotic promoter; brown arrow, therapeutic gene; purple box, polyadenylation signal; green box, recombinase attachment site.

regulatory authorities [5]. The risk assessment for its usage in GM plants (e.g., potato line EH92-527-1) has recently spurred controversial discussion between the European Food Safety Authority (EFSA) and the European Medicines Agency (EMEA) about the therapeutic relevance of kanamycin/neomycin in human medicine. According to the EMEA, aminoglycosides such as kanamycin are currently used in prevention and treatment of critical clinical conditions such as hepatic encephalopathy, multidrug resistant tuberculosis, and other serious invasive bacterial infections, and therefore, their current and potential future use cannot be classified as having a minor therapeutic relevance [EMEA/CVMP/56937/2007]. In this regard, it can be anticipated that the usage of antibiotic resistance markers encoded on plasmids used for animal and human therapy or for the production of biopharmaceuticals will be more critically evaluated by the regulatory authorities in the near future.

Antibiotic resistance genes and metabolic burden

Besides biosafety issues, the use of resistance markers, which are constitutively expressed from the plasmid backbone, also has detrimental consequences for the large-scale production of pDNA for therapeutic purposes. The amplification of plasmids harboring an antibiotic selection marker under conditions mimicking those used at the industrial scale was found to increase the levels of the alarmone ppGpp and the production of inclusion bodies in an *E. coli* host when compared with marker-free variants [26]. In addition, carriage of a therapeutic plasmid induced a significant decrease in growth rate and biomass yield in *E. coli* DH1. Metabolic flux analysis revealed that these effects are due to the increased ATP synthesis requirement arising from constitutive expression of the *nptII* resistance marker, which accounted for 18% of total cell protein [32].

Plasmid size reduction improves structural stability

DNA structural instability can be defined as a series of spontaneous events that culminate in an unforeseen rearrangement, loss, or gain of genetic material. Such events are frequently triggered by the transposition of mobile elements or by the presence of unstable elements such as noncanonical (non-B) structures. Insertion sequences (ISs) are small cryptic mobile elements, ubiquitous in most eubacteria and archae, and tremendously relevant to the field of plasmid bioprocessing. ISs can severely impact plasmid function and yield, by leading to deletions and rearrangements, activation, downregulation or inactivation of neighboring gene expression [33]. Such IS-containing molecules often become predominant throughout the culture stage, whether as a result of high antibiotic selective pressure or the presence of certain genes and regulatory elements capable of enhancing transposition frequency. The *nptII* gene, for example, in particular its 5′ half and approximately 100 bp upstream, constitutes a hotspot for the transposition of several insertion sequences (ISs), including IS*1* [34], IS*2* [35], IS*3* [36], IS*150* [36], and IS*186* [36].

Accessory regions pertaining to the bacterial backbone may also engage in a wide range of structural instability phenomena. Well-known catalysts of genetic instability include direct, inverted, and tandem repeats, which are known to be conspicuous in a large number of commercially available cloning and expression vectors. Such repeats can reach sizes above 0.1 kb, and preferentially map within noncoding regions, SV40 or CMV enhancer/promoter elements, and regulatory sequences (e.g., transcriptional terminators or binding sites) [37]. Over the years, several groups have described the occurrence of spontaneous repeat-mediated slippage events in pDNA, yielding monomeric (deleted) and heterodimeric (amplified) aberrant forms [37–39]. Despite the usual low frequency of each deletion-formation event (typically between 10^{-9} and 10^{-3}) [40], the detrimental impact of these contaminant populations in terms of productivity loss should not be disregarded

[41,42] if one considers the cumulative of all mutational events taking place in a single plasmid. Therefore, the reduction or complete elimination of extraneous noncoding backbone sequences would pointedly reduce the propensity for such events to take place, and consequently, the overall recombinogenic potential of the plasmid.

Antibiotic resistance genes cause gene expression and epigenetic changes in mammalian cells

When placed downstream of a dual prokaryotic and eukaryotic promoter such as PGK-EM7, the *nptII* gene can also be used as a marker for the generation of stable cell lines. Expression of *nptII* in mammalian cells is associated with drawbacks, not yet fully recognized by scientists. However, some interesting publications do exist in the literature. In one example, the *nptII* gene present in plasmids used to transfect fibroblasts was shown to induce a marked decrease in key glycolytic products and a significant downregulation of procollagen 1a and fibronectin genes [43]. Such metabolic and gene expression changes apparently take place via upregulation of the *c-myc* protooncogene through a rather unclear pathway. According to a recent study, deep sequencing performed on several expression vectors has revealed the existence of a complex transcriptional landscape, including significant amounts of unexpected RNA transcripts [44]. Among these, a population of edited sense and antisense small RNAs originating at the *nptII* gene was found capable of inhibiting the expression of co-transfected reporter genes via formation of double-stranded RNA (dsRNA). In addition to these changes, another study demonstrated global epigenetic changes once *nptII* was introduced into mammalian cells [45]. Building on these findings, the authors firmly discourage the use of the *nptII* resistance marker for the generation of stable cell lines.

Bacterial backbone sequences are immunogenic

The immune system of vertebrates has evolved in a way that it is able to recognize bacterial DNA that is rich in unmethylated CpGs. CpGs are known to be statistically underrepresented (~fivefold below expectation) within the eukaryotic genome, assumingly because they mutate at high frequency (~30-fold higher than for other dinucleotides). Throughout the eukaryotic genome, these CG dinucleotides are commonly methylated with the exception of short unmethylated regions, the so-called CpG islands, which can be found within gene regulatory elements. pDNA derived from *E. coli* contains short sequences of nonmethylated CpG dinucleotides in a certain base context [46,47]. Recognition is mediated by Toll-like receptor 9 (TLR-9) in cells of the innate immune system, triggering an inflammatory reaction (activating B cells, monocytes, macrophages, dendritic cells, and natural killer cells) that in turn drives the adaptive immune response towards the vector-encoded gene of interest (GOI). This response is highly appreciated in the case of DNA vaccines and undesirable in the case of gene therapeutic applications where sustained transgene expression is necessary. Therefore, numerous approaches that either lower the immunogenic profile of pDNA vectors by reducing the CpG content [48–51] or increase the immunogenicity through the design of potent immunostimulatory CpG-rich sequences [2,52] have been developed during recent years. For example, in a pUC-based origin of replication and in the *nptII* gene, one finds 47 and 74 CpG sites, respectively (Figure 3). Therefore, using alternative antibiotic resistance-free plasmid selection techniques can alter the immunogenic profile of a plasmid vector. CpG-depleted vectors can be generated by sequence analysis, *in silico* sequence optimization, and DNA synthesis. However, the benefit of complete depletion of CpG motifs, even in the coding sequence of the transgene, is controversial. Recent studies have shown that CpG depletion clearly diminishes the *de novo* transcription of the GOI [53,54]. Another recent study has shown that CpG-free plasmids are taken up more extensively and in contrast are cleared faster from mouse lungs than CpG-rich plasmids [55]. CpG-mediated gene silencing seems to be more pronounced when polyethylenimine (PEI) is used for transfection of pDNA [56,57].

Bacterial backbone sequences induce gene silencing

One of the major obstacles in nonviral gene delivery is the transcriptional silencing of the administered DNA vector in quiescent tissues. The detailed mechanism underlying transgene silencing in mammalian cells still remains elusive due to the complex, dynamic interplay of multiple cellular modification systems. The past 10 years have provided evidence that rapid transgene silencing is due to the formation of inactive chromatin within the episomal vector [58,59]. Chromatin profiling of a DNA minicircle and a conventional plasmid in mouse liver revealed persistent transgene expression with the minicircle whereas the conventional plasmid was silenced. Chromatin immunoprecipitation (ChIP) revealed that silencing of the conventional plasmid was accompanied by an increase in heterochromatin-associated histone modifications and a decrease in modifications typically associated with euchromatin [60]. One common feature of bacterial backbone sequences is that they are transcriptionally inactive in mammalian cells. It is proposed that these transcriptionally inert sequences are similar to the heterochromatin found in eukaryotic cells and are, therefore, responsible for triggering a nucleosome condensation process, finally leading to the silencing of a vicinal transgene [61]. Follow-up studies showed that (i) silencing occurs at a nuclear stage due to chromatin-linked transcriptional blockage, possibly caused by dramatic enrichment of H3K27 trimethylation on plasmid sequences [62] and that (ii) the extragenic spacer length between the 5′ and 3′ end of the eukaryotic expression unit is the major factor influencing transgene silencing [63]. This finding is particularly interesting because it was demonstrated that the source of the spacer sequence is negligible.

Perspectives

Marker-free strategies for other biotechnological applications

Transgenic plants. Agricultural crops are one of the most cited examples of GM organisms, and certainly one of the most contentious. Almost 20 years after the first GM crop was approved for commercialization [64], the controversy over the safety of these products still flares up, often as a

result of unsubstantiated conjectures. Adding to socioeconomic, political, religious, and ethical concerns, GM products have sparked fear of transgene or selectable marker spread to other organisms, as well as of uncontrolled release to the environment. As stated above, the *nptII* gene is among the most frequently used markers for selection of plant cells that carry a desired allele. The essence of the decision in using this gene relied on its ubiquitous presence in the resistome of naturally occurring soil bacteria, and on the belief that the antibiotics to which it confers resistance have an extremely limited use in both human and veterinary clinical settings (see section *Biosafety* for in-depth discussion).

In view of all the above, removal of the selectable marker would substantially simplify the regulatory process and appease divergences, ultimately improving consumer acceptance. This growing need for clean marker-free GM plants has led to strategies that either involve the excision or segregation of these genes from the genome after regeneration of the transgenic plant, or make use of markers that are not based on antibiotic or herbicide resistance. The former may involve co-transformation approaches (integration of marker and transgene on different loci subsequently separated by sexual reproduction), site-specific recombination systems (Cre/*loxP*, flippase/flippase recognition target (FLP/FRT), or R/RS), homologous recombination, or transposition systems. Positive and negative selection systems relying on the growth of plant cells in the presence of unusual nutrients and other factors have also been exploited. Thorough reviews on this matter can be found elsewhere [65].

Gene delivery to stem cells

Stem cells hold great promise for biomedical research and treatment of several incapacitating or incurable conditions. For example, when combined with gene delivery, they have the potential to enhance tissue regeneration or to serve as vehicles for delivery of therapeutic proteins. The tantalizing possibility of treating genetic diseases by autologous transplantation of genetically corrected patient-specific stem cells became even more real after the groundbreaking achievements in the derivation of induced pluripotent stem cells from somatic cells [66]. Nevertheless, realization of the abovementioned potential very much depends on developing efficient and optimized strategies for genetic manipulation of stem cells. Traditional cocktail delivery approaches rely on randomly integrating retroviral or lentiviral vectors, but notwithstanding, the latter have raised concerns about the possible activation of oncogenes, disruption of tumor suppressor genes, and generation of genomic heterogeneity. In the past few years, plasmid-based vectors were proposed as safer nonviral and nonintegrating reprogramming alternatives, despite showing significantly lower reprogramming efficiencies than viral-based methods (typically ranging from 0.01 to 0.5%) [67]. Particularly, plasmid minicircles were shown to successfully reprogram human adipose-derived stem cells in a 2–4 week period, yielding higher overall efficiencies than those seen with their standard plasmid counterparts (0.005% and 0.001%, respectively), and with no detectable evidence of genome integration events [68,69]. In another study, it was observed that size-reduced (35% smaller) marker-free plasmids were capable of efficiently transfect human mesenchymal and mouse neural stem cells, achieving 2.6-fold higher intracellular copy numbers and twice the number of fluorescent cells than those observed with the parental plasmid [27]. A follow-up of this study using minicircles further substantiated the benefits in terms of higher cell viability and transgene expression [70]. Similar results were obtained more recently with GFP-encoded minicircles electrotransferred both *in vitro* and *in vivo* to mouse melanoma cells [71].

Recombinant protein expression

To date, antibiotics are commonly used during bacterial fermentation or transient protein production in mammalian cells. The last several years have seen the development of various antibiotic resistance-free plasmid vector systems that can be utilized for recombinant protein production in *E. coli*. These systems either rely on toxin/antitoxin post-segregational killing [72] or essential gene complementation [73,74]. In addition to the improved biosafety profile, some of these systems have been able to substantially increase final recombinant protein yields. Upon functional validation on an industrial scale it is very likely that regulatory authorities will soon request these alternative selection means.

For mammalian cell technology, alternatives to positive selection using antibiotics or cytotoxic drugs are rare, although the usage of antibiotic selection has been repeatedly associated with drawbacks [75,76]. A recent, interesting approach has combined the sleeping beauty transposon system encoded on a minicircle, with chemically-induced growth factor dimerization selection [77]. Concerning transient protein expression in mammalian cells it is very likely that smaller, marker-free plasmids will improve transformation efficacy (for details, see Box 1) and consequently final product titers, although to the best of our knowledge their usage has not yet been reported.

Concluding remarks and future trends

Although seemingly convenient, the use of antibiotic selection relies on a 40-year-old approach of genetic construction that dates back to the founding years of recombinant DNA technology. In this review, we have focused on the state-of the-art achievements in marker-free technology and revisited the motivations that have secured their successful implementation. Although still faced with considerable challenges (Box 2), more is yet to be expected from these systems. It is conceivable that emergent fields such as that of synthetic biology, for example, may in the near future largely benefit from marker-free approaches. In fact, when compared to genome editing, plasmids offer several advantages for synthetic biology applications, particularly due to their modular structure, easy manipulation, and large gene dosage capacity (if present in multiple copies). In this regard, one can thus expect that trans-acting dependency devices will be preferred over antibiotic resistance markers as a biological confinement strategy to curb genetic pollution (recently reviewed in [78]). The next few years are likely to witness exciting progress in marker-free technologies, which will unfold in tandem with an

Box 2. Outstanding questions

- Close evaluation of standardized plasmid backbone sequences will be necessary to gain understanding at the systems level; the question of how the configuration of different elements on the plasmid backbone influences the final performance in mammalian cells remains elusive. What is the influence of antibiotic resistance genes on plasmid performance in mammalian cells?
- Detailed characterization of the impact of unwanted spurious transcripts and aberrant forms that coexist with a parental plasmid on performance and safety is still missing. What is the difference between mammalian cells containing conventional and marker-free plasmids when investigated using multiple '-omics' methods?
- Comparative studies benchmarking the different marker-free approaches are currently missing. How do these approaches differ in overall performance? How do they differ in manufacturing costs?
- Detailed investigation of the geographical and ecological distribution of synthetic plasmid variants should be performed. Are synthetic plasmids that spread to the environment really an unrecognized source of antibiotic resistance?

increased perception of the extent to which these systems may be applied across different research fields.

References

1 Azzoni, A.R. *et al.* (2007) The impact of polyadenylation signals on plasmid nuclease-resistance and transgene expression. *J. Gene Med.* 9, 392–402
2 Coban, C. *et al.* (2005) Effect of plasmid backbone modification by different human CpG motifs on the immunogenicity of DNA vaccine vectors. *J. Leukoc. Biol.* 78, 647–655
3 Uchijima, M. *et al.* (1998) Optimization of codon usage of plasmid DNA vaccine is required for the effective MHC class I-restricted T cell responses against an intracellular bacterium. *J. Immunol.* 161, 5594–5599
4 Mairhofer, J. *et al.* (2008) A novel antibiotic free plasmid selection system: advances in safe and efficient DNA therapy. *Biotechnol. J.* 3, 83–89
5 FDA (1996) Points to consider on plasmid DNA vaccines for preventive infectious disease indications. *Cent. Biol. Eval. Res.* Docket No. 96N-0400
6 Darquet, A.M. *et al.* (1997) A new DNA vehicle for nonviral gene delivery: supercoiled minicircle. *Gene Ther.* 4, 1341–1349
7 Bigger, B.W. *et al.* (2001) An araC-controlled bacterial cre expression system to produce DNA minicircle vectors for nuclear and mitochondrial gene therapy. *J. Biol. Chem.* 276, 23018–23027
8 Chen, Z.Y. *et al.* (2003) Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo. *Mol. Ther.* 8, 495–500
9 Jechlinger, W. *et al.* (2004) Minicircle DNA immobilized in bacterial ghosts: in vivo production of safe non-viral DNA delivery vehicles. *J. Mol. Microbiol. Biotechnol.* 8, 222–231
10 Schleef, M. (2013) *Minicircle and Miniplasmid DNA Vectors: The Future of Non-Viral and Viral Gene Transfer*. Wiley-VCH
11 Soubrier, F. *et al.* (1999) pCOR: a new design of plasmid vectors for nonviral gene therapy. *Gene Ther.* 6, 1482–1488
12 Soubrier, F. *et al.* (2005) Improvement of pCOR plasmid copy number for pharmaceutical applications. *Appl. Microbiol. Biotechnol.* 66, 683–688
13 Fabre, E.E. *et al.* (2008) Careful adjustment of Epo non-viral gene therapy for beta-thalassemic anaemia treatment. *Genet. Vaccines Ther.* 6, 10
14 Nikol, S. *et al.* (2008) Therapeutic angiogenesis with intramuscular NV1FGF improves amputation-free survival in patients with critical limb ischemia. *Mol. Ther.* 16, 972–978
15 Maulik, N. (2009) NV1FGF, a pCOR plasmid-based angiogenic gene therapy for the treatment of intermittent claudication and critical limb ischemia. *Curr. Opin. Investig. Drugs* 10, 259–268
16 Marie, C. *et al.* (2010) pFARs, plasmids free of antibiotic resistance markers, display high-level transgene expression in muscle, skin and tumour cells. *J. Gene Med.* 12, 323–332
17 Williams, S.G. *et al.* (1998) Repressor titration: a novel system for selection and stable maintenance of recombinant plasmids. *Nucleic Acids Res.* 26, 2120–2124
18 Cranenburgh, R.M. *et al.* (2001) *Escherichia coli* strains that allow antibiotic-free plasmid selection and maintenance by repressor titration. *Nucleic Acids Res.* 29, E26
19 Guimaraes-Walker, A. *et al.* (2008) Lessons from IAVI-006, a phase I clinical trial to evaluate the safety and immunogenicity of the pTHr.HIVA DNA and MVA.HIVA vaccines in a prime-boost strategy to induce HIV-1 specific T-cell responses in healthy volunteers. *Vaccine* 26, 6671–6677
20 Jaoko, W. *et al.* (2008) Safety and immunogenicity of recombinant low-dosage HIV-1 A vaccine candidates vectored by plasmid pTHr DNA or modified vaccinia virus Ankara (MVA) in humans in East Africa. *Vaccine* 26, 2788–2795
21 McCormack, S. *et al.* (2008) EV02: a Phase I trial to compare the safety and immunogenicity of HIV DNA-C prime-NYVAC-C boost to NYVAC-C alone. *Vaccine* 26, 3162–3174
22 Garmory, H.S. *et al.* (2005) Antibiotic-free plasmid stabilization by operator-repressor titration for vaccine delivery by using live Salmonella enterica Serovar typhimurium. *Infect. Immun.* 73, 2005–2011
23 Leckenby, M.W. *et al.* (2009) Enhanced vaccine antigen delivery by Salmonella using antibiotic-free operator-repressor titration-based plasmid stabilisation compared to chromosomal integration. *Microb. Pathog.* 46, 201–206
24 Huang, J.M. *et al.* (2010) Oral delivery of a DNA vaccine against tuberculosis using operator-repressor titration in a Salmonella enterica vector. *Vaccine* 28, 7523–7528
25 Luke, J. *et al.* (2009) Improved antibiotic-free DNA vaccine vectors utilizing a novel RNA based plasmid selection system. *Vaccine* 27, 6454–6459
26 Mairhofer, J. *et al.* (2010) Marker-free plasmids for gene therapeutic applications–lack of antibiotic resistance gene substantially improves the manufacturing process. *J. Biotechnol.* 146, 130–137
27 Ribeiro, S. *et al.* (2012) Plasmid DNA size does affect nonviral gene delivery efficiency in stem cells. *Cell. Reprogram.* 14, 130–137
28 Lu, J. *et al.* (2013) A mini-intronic plasmid (MIP): a novel robust transgene expression vector in vivo and in vitro. *Mol. Ther.* 21, 954–963
29 Furger, A. *et al.* (2002) Promoter proximal splice sites enhance transcription. *Genes Dev.* 16, 2792–2799
30 Wellington, E.M. *et al.* (2013) The role of the natural environment in the emergence of antibiotic resistance in gram-negative bacteria. *Lancet Infect. Dis.* 13, 155–165
31 Chen, J. *et al.* (2012) A survey of drug resistance bla genes originating from synthetic plasmid vectors in six Chinese rivers. *Environ. Sci. Technol.* 46, 13448–13454
32 Rozkov, A. *et al.* (2004) Characterization of the metabolic burden on *Escherichia coli* DH1 cells imposed by the presence of a plasmid containing a gene therapy sequence. *Biotechnol. Bioeng.* 88, 909–915
33 Mahillon, J. and Chandler, M. (1998) Insertion sequences. *Microbiol. Mol. Biol. Rev.* 62, 725–774
34 Prather, K.L. *et al.* (2006) Identification and characterization of IS1 transposition in plasmid amplification mutants of *E. coli* clones producing DNA vaccines. *Appl. Microbiol. Biotechnol.* 73, 815–826
35 Oliveira, P.H. *et al.* (2009) Deletion formation mutations in plasmid expression vectors are unfavored by runaway amplification conditions and differentially selected under kanamycin stress. *J. Biotechnol.* 143, 231–238
36 Szeverenyi, I. *et al.* (1996) Vector for IS element entrapment and functional characterization based on turning on expression of distal promoterless genes. *Gene* 174, 103–110
37 Oliveira, P.H. *et al.* (2010) Analysis of DNA repeats in bacterial plasmids reveals the potential for recurrent instability events. *Appl. Microbiol. Biotechnol.* 87, 2157–2167
38 Ribeiro, S.C. *et al.* (2008) High frequency plasmid recombination mediated by 28 bp direct repeats. *Mol. Biotechnol.* 40, 252–260
39 Bi, X. and Liu, L.F. (1996) A replicational model for DNA recombination between direct repeats. *J. Mol. Biol.* 256, 849–858
40 Oliveira, P.H. *et al.* (2008) Recombination frequency in plasmid DNA containing direct repeats–predictive correlation with repeat and intervening sequence length. *Plasmid* 60, 159–165

41 Borgi, I. and Gargouri, A. (2008) A spontaneous direct repeat deletion in the pGEX fusion vector decreases the expression level of recombinant proteins in *Escherichia coli*. *Protein. Expr. Purif.* 60, 15–19
42 Kim, M. *et al.* (2011) A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies. *Biotechnol. Bioeng.* 108, 2434–2446
43 Valera, A. *et al.* (1994) Expression of the neomycin-resistance (neo) gene induces alterations in gene expression and metabolism. *Hum. Gene Ther.* 5, 449–456
44 Nejepinska, J. *et al.* (2012) Deep sequencing reveals complex spurious transcription from transiently transfected plasmids. *PLoS ONE* 7, e43283
45 Dutta, P. *et al.* (2012) Global epigenetic changes induced by SWI2/SNF2 inhibitors characterize neomycin-resistant mammalian cells. *PLoS ONE* 7, e49822
46 Yamamoto, S. *et al.* (1992) Unique palindromic sequences in synthetic oligonucleotides are required to induce IFN [correction of INF] and augment IFN-mediated [correction of INF] natural killer activity. *J. Immunol.* 148, 4072–4076
47 Krieg, A.M. *et al.* (1995) CpG motifs in bacterial DNA trigger direct B-cell activation. *Nature* 374, 546–549
48 Yew, N. *et al.* (2000) Reduced inflammatory response to plasmid DNA vectors by elimination and inhibition of immunostimulatory CpG motifs. *Mol. Ther.* 1, 255–262
49 Yew, N. *et al.* (2002) CpG-depleted plasmid DNA vectors with enhanced safety and long-term gene expression in vivo. *Mol. Ther.* 5, 731–738
50 Hyde, S.C. *et al.* (2008) CpG-free plasmids confer reduced inflammation and sustained pulmonary gene expression. *Nat. Biotechnol.* 26, 549–551
51 Navarro, G. *et al.* (2010) Low generation PAMAM dendrimer and CpG free plasmids allow targeted and extended transgene expression in tumors after systemic delivery. *J. Control. Release* 146, 99–105
52 Higgins, D. *et al.* (2007) Immunostimulatory DNA as a vaccine adjuvant. *Expert Rev. Vaccines* 6, 747–759
53 Kosovac, D. *et al.* (2010) Minimal doses of a sequence-optimized transgene mediate high-level and long-term EPO expression in vivo: challenging CpG-free gene design. *Gene Ther.* 18, 189–198
54 Bauer, A.P. *et al.* (2010) The impact of intragenic CpG content on gene expression. *Nucleic Acids Res.* 38, 3891–3908
55 Lesina, E. *et al.* (2010) CpG-free plasmid DNA prevents deterioration of pulmonary function in mice. *Eur. J. Pharm. Biopharm.* 74, 427–434
56 Lesina, E. *et al.* (2010) The effect of CpG motifs on gene expression and clearance kinetics of aerosol administered polyethylenimine (PEI)-plasmid DNA complexes in the lung. *J. Control. Release* 143, 243–250
57 Davies, L.A. *et al.* (2012) The use of CpG-free plasmids to mediate persistent gene expression following repeated aerosol delivery of pDNA/PEI complexes. *Biomaterials* 33, 5618–5627
58 Suzuki, M. *et al.* (2006) Plasmid DNA sequences present in conventional herpes simplex virus amplicon vectors cause rapid transgene silencing by forming inactive chromatin. *J. Virol.* 80, 3293–3300
59 Riu, E. *et al.* (2007) Histone modifications are associated with the persistence or silencing of vector-mediated transgene expression in vivo. *Mol. Ther.* 15, 1348–1355
60 Riu, E. *et al.* (2005) Increased maintenance and persistence of transgenes by excision of expression cassettes from plasmid sequences in vivo. *Hum. Gene Ther.* 16, 558–570
61 Chen, Z. *et al.* (2004) Silencing of episomal transgene expression by plasmid bacterial DNA elements in vivo. *Gene Ther.* 11, 856–864
62 Gracey Maniar, L.E. *et al.* (2013) Minicircle DNA vectors achieve sustained expression reflected by active chromatin and transcriptional level. *Mol. Ther.* 21, 131–138
63 Lu, J. *et al.* (2012) The extragenic spacer length between the 5′ and 3′ ends of the transgene expression cassette affects transgene silencing from plasmid-based vectors. *Mol. Ther.* 20, 2111–2119
64 Marineau, B. (2001) *First Fruit: the Creation of the Flavr Savr Tomato and the Birth of Genetically Engineered Food*. McGraw-Hill
65 Darbani, B. *et al.* (2007) Methods to produce marker-free transgenic plants. *Biotechnol. J.* 2, 83–90
66 Takahashi, K. and Yamanaka, S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676
67 Si-Tayeb, K. *et al.* (2010) Generation of human induced pluripotent stem cells by simple transient transfection of plasmid DNA encoding reprogramming factors. *BMC Dev. Biol.* 10, 81
68 Jia, F. *et al.* (2010) A nonviral minicircle vector for deriving human iPS cells. *Nat. Methods* 7, 197–199
69 Narsinh, K.H. *et al.* (2011) Generation of adult human induced pluripotent stem cells using nonviral minicircle DNA vectors. *Nat. Protoc.* 6, 78–88
70 Madeira, C. *et al.* (2013) Nonviral gene delivery to neural stem cells with minicircles by microporation. *Biomacromolecules* 13, 1379–1387
71 Chabot, S. *et al.* (2012) Minicircle DNA electrotransfer for efficient tissue-targeted gene delivery. *Gene Ther.* 20, 62–68
72 Peubez, I. *et al.* (2010) Antibiotic-free selection in *E. coli*: new considerations for optimal design and improved production. *Microb. Cell Fact.* 9, 65
73 Vidal, L. *et al.* (2008) Development of an antibiotic-free plasmid selection system based on glycine auxotrophy for recombinant protein overproduction in *Escherichia coli*. *J. Biotechnol.* 134, 127–136
74 Dong, W.R. *et al.* (2010) Novel antibiotic-free plasmid selection system based on complementation of host auxotrophy in the NAD de novo synthesis pathway. *Appl. Environ. Microbiol.* 76, 2295–2303
75 Kaufman, W.L. *et al.* (2008) Homogeneity and persistence of transgene expression by omitting antibiotic selection in cell line isolation. *Nucleic Acids Res.* 36, e111
76 Lanza, A.M. *et al.* (2013) Evaluating the influence of selection markers on obtaining selected pools and stable cell lines in human cells. *Biotechnol. J.* 8, 811–821
77 Kacherovsky, N. *et al.* (2012) Combination of Sleeping Beauty transposition and chemically induced dimerization selection for robust production of engineered cells. *Nucleic Acids Res.* 40, e85
78 Wright, O. *et al.* (2013) Building-in biosafety for synthetic biology. *Microbiology* 159, 1221–1235
79 Lewis, L.A. *et al.* (2012) Protein-DNA interactions define the mechanistic aspects of circle formation and insertion reactions in IS2 transposition. *Mobile DNA* 3, 1
80 Ribeiro, S.C. *et al.* (2004) The role of polyadenylation signal secondary structures on the resistance of plasmid vectors to nucleases. *J. Gene Med.* 6, 565–573
81 Lukacs, G.L. *et al.* (2000) Size-dependent DNA mobility in cytoplasm and nucleus. *J. Biol. Chem.* 275, 1625–1629
82 Kreiss, P. *et al.* (1999) Plasmid DNA size does not affect the physicochemical properties of lipoplexes but modulates gene transfer efficiency. *Nucleic Acids Res.* 27, 3792–3798
83 Yin, W. *et al.* (2005) Investigations of the effect of DNA size in transient transfection assay using dual luciferase system. *Anal. Biochem.* 346, 289–294