**CANDIDATE PHAGES TO DELIVER THE HIJACK, DETECT AND TERMINATE SYSTEM**

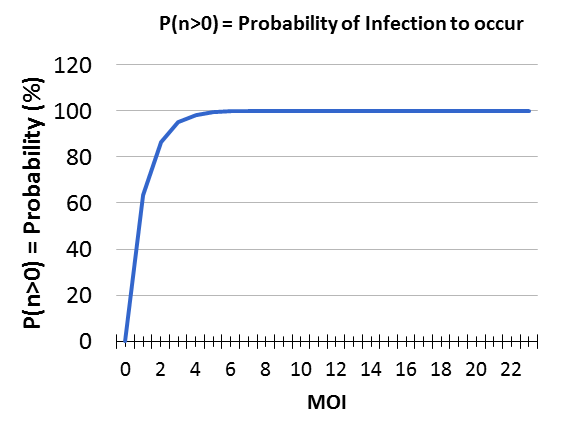
As a proof of concept for the expression in *Mycobacteria spp* of the Hijack, Detect and Terminate (HDT) genetic system it is interesting to test the synthetic construct in *Escherichia coli* by using Enterobacteria phages and also to find a way to do the same in *Mycobacteria spp* using Mycobacteriophages.

A common approach for the expression of synthetic devices in these hosts is the use of conventional protocols for the chemical transformation or electro-transformation of competent cells. Nevertheless, a delivery system based on these transformation protocols is not suitable on the conditions for which this HDT system has been conceived, since this is meant to be targeting extremely low concentrations of target host (i.e. *M. tuberculosis*) with the final goal to speed up its growth rate in a few hours through the Hijack module so that it can be easily detected with the integrated Detection module. All of it in the most cost-effective manner, since the detection device is should be able to reach the most number of users in hospitals or in field applications.

Therefore, it is important to dispose of a way in which it is possible to transform very low amounts of the target strain with high transformation efficiencies. If we are to rely on chemical transformation for the delivery of this HDT circuit into the target strain it is evident that the overall cost of this detection device will be hampered by the transformation efficiency, which is no larger than 2.7·105 number of transformants per µg of DNA and per 109 electroporated cells for *M. tuberculosis* and 7.4·102 for *M. smegmatis* at 37 º, see **Table 1** below (Wards, B. J., & Collins, D. M., 1996).

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 1**. Electroporation of the relevant slow-growing mycobacteria strains with pYUB18 at different temperatures. Modified and adapted from Wards, B. J., & Collins, D. M., 1996) | | | |
| Strain | Efficiency (Number of transformants per µg DNA per 109 cells electroporated) | | |
| 0ºC | 20ºC | 37ºC |
| M. tuberculosis H37Rv | 8.5 x 103 | 1.2 x 105 | 2.7 x 105 |
| M. smegmatis mc2 155 | 3.4 x 104 | 4.4 x 103 | 7.4 x 102 |

Furthermore, designing a platform or device for the rapid and easy chemical or electro-transformation of highly heterogeneous samples with very low concentrations of target strains can suppose the development of a cost-ineffective, difficult to transport and time consuming system based on the actual procedures for transformation. Therefore the use of the conditionally replicating phages designed here present themselves as an alternative with many advantages such as the use of the viral cycle for the massive production of the HDT devices which is to be stored under the protection of a bacteriophage capsid. This allows for long term storage of quality DNA and is also the most efficient manner in which to transform target cells with the highest efficiency, which is regulated with the proportion between the phages used against bacteria. Therefore, by using values of the Multiplicity of Infection (MOI) higher than 1, see **Figure 1**, there is a good chance that the minimum amount of target strain present in a target sample will be infected, an therefore transduced to express the HDT synthetic DNA construct. In the following Figure 1, it is plotted the probability of infection according to the proportion of bacteriophages respective of its host following the Poisson distribution.



**Figure 1**. Probability of infection to occur following a Poisson distribution. **P(n>0) :** probability of infection when the number “n” of virions is greater than n=0. **K :** number of virions. **MOI:** Multiplicity of infection.

Therefore, it can be stated that using bacteriophage in appropriate amounts as a method for the delivery of the HDT system provides the following advantages:

1. Ability to transform (transduce) very low amounts of target bacteria
2. Higher transformation efficiencies
3. Long term storage of quality DNA
4. Mass production of synthetic bacteriophages harbouring the HDT synthetic DNA.

Hence, it is interesting to design bacteriophages that can deliver the HDT synthetic genetic circuit and to perform these designs taking into account the advantages and disadvantages of the viral cycle, being its lytic nature and the capability to host a large size of synthetic DNA the main two concerns. Towards the pursuing of this goal, two suitable designs, one for its expression in *Escherichia coli* and the other for *Mycobacterium spp* have been exposed here.

In regard to Enterobacteria phages, although the different classes of these phages are very well characterised, there is a lack of phages with deleted non-essential regions that can host large synthetic genetic constructs such as the one designed here ( ~9000 bp). And although bacteriophages are able to pack genomes greater than its wild-type length, there are limitations. The **Table 2** summarizes package limits for common *Escherichia coli* bacteriophages and for the TM4 mycobacteriophage.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 2**. DNA Cutting and Packaging in phages (adapted form Knipe, D. M., et al 2007) | | | | |
| Phage | Packaging substrate | First cut | Second Cut | Packaging Limits (% Genome) |
| T4 | Concatamer | Random | Headful | 102 |
| P22 | Concatamer | pac site | Headful | 109 |
| ʎ | Concatamer | cos site | cos site | 79-110 |
| T7 | Concatamer | Terminal overlap | Terminal repeat | 85-103 |
| ΦX174 | Dimer ss rolling circle | Replication origin | Replication origin | Limited |
| M13 | Dimer ss rolling circle | Replication origin | Replication origin | Unlimited |
| Mu-1 | Inserted monomer | Near left end | Headful | 105 |
| P1 | Concatamer | pac site | Headful | 110 |
| TM4 / phAE159 | Unknown | Unknown | Unknown | Unknown |

The Table 2 shows that T7 and T4 and ʎ are good candidates for hosting additional synthetic genetic constructs since they can pack genomes larger than its wild-type length, but the most suitable ones here are the ʎ and T4 since they have larger genomes able to host larger % excess.

The main issue here is that the genetic constructs that are to be tested in the phage are ~9000 bp and it makes not possible to use T4 or ʎ as a host for these constructs unless the genomes of these phages are modified to become genome-minimized bacteriophages.

Minimizing the genome to host synthetic DNA is a similar approach used for the TM4-like phAE159 mycobacteriophage proposed here (see Jacobs Jr, W. R. 2014). Therefore, it should be possible to build a version of the T4 or the ʎ phage that lacks most of their non-essential genes to make space for synthetic constructs. The phAE159 is an essential version of the TM4 mycobacteriophage that contains a deleted region corresponding to non-essential genes. This region is substituted by a shuttle construct flanked by *PacI* restriction sites where an exchangeable insert containing genetic components for the packaging in a ʎ capsid and the replication in *Escherichia coli* is present (click ʎ-like ʎgt11 and TM4-like phAE159 below)

This shuttle plasmid (rather a shuttle bacteriophage vector), once *in vitro* packaged in ʎ capsids (since it complies with the minimum genome length for packaging in ʎ) (see Table 2), will be able to replicate in Escherichia coli as a plasmid, and in Mycobacteria as a Mycobacteriophage, since it contains all the essential genes that give identity to a Mycobacteriophage.

Here we porpose to use a commercial version of the ʎ phage (ʎgt11) for *Escherichia coli* and the available phAE159 TM4-like mycobacteriophage for *Mycobacterium spp* as delivery channels for the HDT. The T4 enterobacteria phage has been discarded because generating enough available space for incorporating the HDT synthetic device is considered highly costly, given the number of genes that would need to be deleted.

The ʎ bacteriophage wild-type genome contains 48502 bp and is able to host up to 110 % of its genome (Knipe, D. M., et al. 2007). There are in the market commercially available versions of the lambda bacteriophage for its use in building DNA libraries and expression vectors. One of these commercially available vectors is the ʎgt11 which is a 43700 Kb expression vector containing a unique EcoRI restriction site within a lacZ gene positioned 53 bp upstream from the termination codon. Open reading frames within the inserted DNA fragments may be expressed as fusion proteins with the lacZ gene product, β-galactosidase allowing to screen for recombinants using the blue-white screening in presence of IPTG and Xgal. The ʎgt11 contains a mutation in the *cI* gene (i.e. *cIts857*) that makes it possible to be switched between lysogen or lytic phage. With the *cIts857* genotype, when infection occurs at 32º the phage lysogenises (i.e. remains in the genome in form of a prophage), but when temperature is switched to 42º the lytic cycle of the phage activates. In this case, this lytic character is of interest since it is a model for the lytic mycobacteriophage TM4.

It is known from the literature that there is a requirement of a minimum and maximum genome length (38 and 53kb, respectively) for the efficient packaging and for the production of viable ʎ phage particles (Chauthaiwale, V.M., et al. 1992). The viability of the bacteriophage decreases when its genome length is greater than 105% or less than 78% of that of wild-type ʎ.

Therefore, in order to design an appropriate bacteriophage, the vector ʎgt11 is here proposed as ideal, since its small size allows provides enough available space to pack between 7000-9000 bp (being 9625 bp its maximum) into the capsids without compromising much its viability. This is enough space for introducing the Hijack, Detection and Termination (HDT) synthetic DNA into the bacteriophage genome to be used as a vector for *Escherichia coli*.

Also there is the analogy between ʎgt11 with the phAE159 version of TM4 mycobacteriophage shown in Jacobs Jr, W. R. (2014) in regard to both containing a large space for the insertion of the synthetic device. In phAE159 is as simple as using the unique dual PacI restriction sites to switch the lambda cassette within phAE159 and any synthetic DNA. In ʎgt11, the insertion site is provided within a unique EcoRI restriction site within the LacZ gene, allowing the screening of recombinants by using the blue/white lacZ test (in presence of XGal and IPTG) in a bacterial host lacking LacZ.

A potential issue may be that phAE159 version of the TM4 bacteriophage is a strict lytic phage, whereas the lambda is a temperate one. Temperate bacteriophages are both lytic and lysogenic, and they are genetically regulated to switch between these two states. Nevertheless, ʎgt11 also contains a mutation in *cI* (i.e. *cI857*) which allows the activation of the lytic cycle when temperature is raised to 42º. Therefore, ʎgt11 becomes a good model of the lytic TM4 (phAE159) mycobacteriophage.

Plasmid systems are available for packaging and expression in *Escherichia coli*, but unfortunately, the same technology is not as straightforward in *Mycobacteria spp* as it is in *Escherichia coli*. As it is explained in Jacobs Jr, W.R. (2014) an interesting approach was to find plasmids compatible between both *Escherichia coli* and *Mycobacterium spp* but this approach is restricted in *Mycobacterium spp* because of one of these reasons: “*(i) the inability of naked DNA to enter the mycobacterial cell, (ii) degradation by a restriction system of the E. coli DNA seen as foreign, (iii) failure to have a functional selection system (appropriate concentrations of antibiotic and sufficiently high levels of expression of the selectable marker gene), and/or (iv) the inability of the mycobacterial plasmid to replicate and segregate into M. smegmatis cells.*” Nonetheless, when mycobacteriophage vectors that include a shuttle sequence for package and replication in lambda capsids, the shuttle vector appears to work in both hosts.

Because of that, the construction of a shuttle phasmid that behaves as plasmid in *Escherichia coli* but as a phage in mycobacteriophage has been a breakthrough for the development of gene transfer systems in interesting species of mycobacterium such as *Mycobacterium tuberculosis*. Furthermore, the capability of having a shuttle vector able to be replicated in *Escherichia coli* enables for unlimited genetic engineering possibilities. This is another reason why plasmids cannot been considered here as a model for testing the delivery of the Hijack, Detect and Terminate device into bacterial hosts, since bacteriophages are a requirement, and the ʎ bacteriophage has been considered instead for its use in *Escherichia coli* as a model of what can be possible in Mycobacteria with the TM4 Mycobacteriophage.

It is important to mention that although the use of plasmids is more advantageous in most cases for the purpose of gene expression, because its stability and the lack of lysis, the use of phages as gene delivery systems is useful in the sense that they can provide high copies of the same construct (this is especially interesting if you are able to inhibit the lytic cycle by building conditionally replicating strains), the high efficiency of introducing recombinant by in vitro packaging followed by infection, and the easiness of screening bacteriophage that give rise to plaques which can be screened by using DNA probes or antibody probes for expression markers (Jendrisak, J., et al. 1987).

In the following Table 3, you have a summarizing comparison of the characteristics of both ʎgt11 and phAE159

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 3**. The table summarizes the gene expression profile possibilities of each of the synthetic phages together with its requirements for each profile. | | | |
|  | Pseudolysogen/Lysogenic  (1 copy/cell) | High copy without lysis | Lytic (high copy) |
| ʎgt11Q-am44 | (30-32º) | At 42º + supE44 | At 42º + supE44 |
| ∆Lys-phAE159 | (37º) | 30º + Non-complementing host | 30º + LysA-LysB-holin complementing host |

**SELECT HERE THE TWO DIFFERENT DESIGNS: Add Two Hyperlinks that drive two different sites:**

ʎ-like ʎgt11 TM4-like phAE159

**REFERENCES**

Wards, B. J., & Collins, D. M. (1996). Electroporation at elevated temperatures substantially improves transformation efficiency of slow‐growing mycobacteria. *FEMS microbiology letters*, *145*(1), 101-105.

**ENGINEERING ʎ-like ʎgt11 PHAGE TO HIJACK, DETECT AND TERMINATE *ESCHERICHIA COLI***

A main characteristic of the ʎgt11 genome is that it contains an amber mutation in the “S” lysis gene (see Table 1). Similar than for TM4-phAE159 it also contains a temperature sensitive repressor (here in form of a cI-ts857 mutation), unknown in phAE159. Although this phage vector has this amber mutation in the gene S required for lysis, it has been demonstrated that the lysis cassette in ʎ bacteriophages is dependent on R, Rz, and S genes (Garett, J., et al. 1981). And this amber mutant has been shown to be able to replicate at 2h post-induction at 42º in non-amber-suppressive strains of *Escherichia coli*. Therefore, this phage is not really a conditionally replicating phage since it disposes of additional genes involved on the lysis. Nevertheless it is important to recognise the contribution of the S gene product to the lysis of the host cell, since when this ʎgt11 is replicated in supF amber suppressor strain, it is able to replicate immediately (without the 2h delay) (see Lin, C. S., et al.1998; Gelvin, S. 2012).

The most interesting part of the Bacteriophage ʎgt11 is that it is commercially available and can be modified to generate an additional mutation in the Q gene, responsible for late expression (i.e. capsid formation, packaging and lysis). Following this strategy, and taking into account that ʎgt11 is commercially available, the generation of an amber mutation (TAG) in the Q gene that could be complemented by a strain of *Escherichia coli* providing the right aminoacid for this phage, would provide a very good candidate for a conditionally replicating bacteriophage for building the Hijack, Detect and Terminate (HDT) Synthetic device developed by the IISER Pune iGEM team.

Lin, C. S., et al.1998 innovated the engineering of ʎgt11 by introducing first an amber mutation in the gene Q and demonstrating that this mutation can be recovered in suitable supE amber suppressor *Escherichia coli* strains such as JM103. When testing the amber mutant phage for the gene Q in a non-amber-supressor strain, the cells did not lyse during the 8h period after the lysis inducing temperature shift (because of *cIts857* mutation) to 42º. That does mean that a time frame of more than 8h is available for the HDT system to promote growth and enable detection.

Here it is proposed to use a similar strategy but by providing a manner in which to build a similar phage by using the last CRISPR/Cas9 technology, in which precise Amber mutations can be introduced intelligently in the phage.

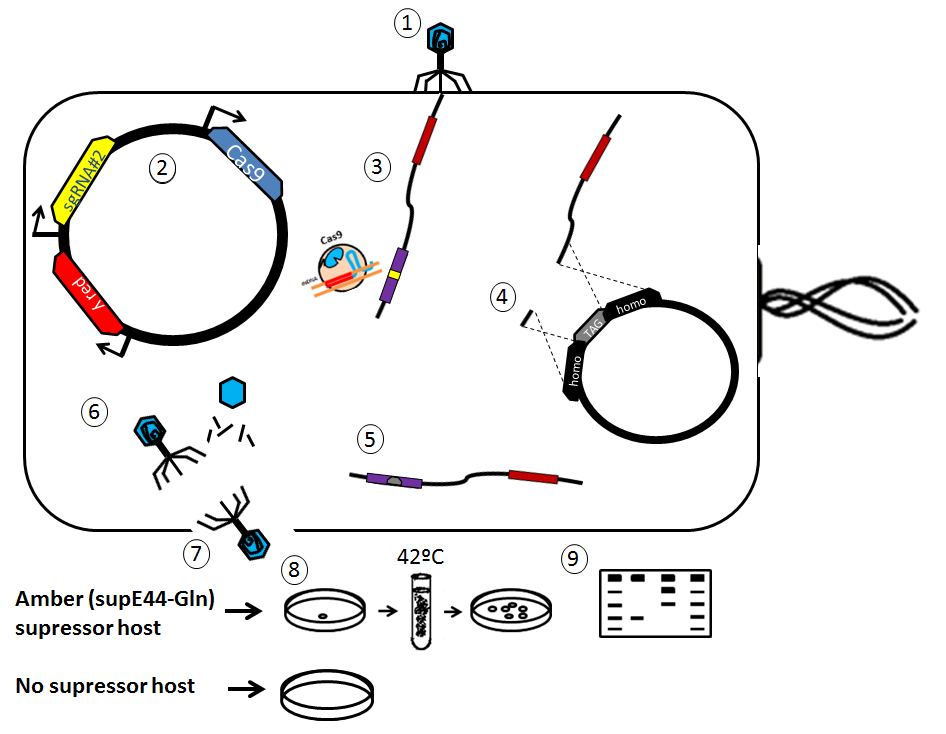
A main issue of playing with ʎgt11 as a model is that there is no complete sequence for this version of the ʎ genome other than the insert region containing the LacZ gene that includes the unique EcoRI restriction site (Moran, L. S., et al. 1990; GenBank: X16973.1), although there can be found many versions of this vector in the literature. However, and according to the genes of our interest, it is right to assume that since ʎgt11, although simplified, is a functional version of the complete wild-type ʎ phage and both are enabled to complete a lytic cycle, it is then possible to find the sequence of the Q gene in the available genome sequence of the ʎ bacteriophage (GenBank: J02459.1).

Thus, the strategy to follow in order to build a lytic phage knocked out for the gene Q, which is responsible of the late gene expression (both lysis and packaging) with the CRISPR/Cas9 system will be as follows below.

The **Table 1**, summarizes the main strains of ʎ bacteriophages mentioned in this protocol, together with the strains of *Escherichia coli* to be used. It also includes the relevant phenotypes to be considered and the expected interactions of each bacteriophage with its respective host strain.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 1.** E. coli strains and bacteriophages versions relevant for this protocol. Notice that ʎgt11/Q-am44 is presented with different, but equivalent, nomenclatures. | | | |
| ***E. coli* stain** |  | **Genotype** | **Relevant Phenotype** |
|  | JM103**(a)** | ∆lac pro thi strA supE endA sbcB15 hsdR4 F’ traD36 proAB lacIq Z∆M15] | Amber suppressor (GlnV44 / supE44) |
|  | DH5α | supE44 ∆(LacZ)M15 recA1 endA1 hsdR17 | Amber suppressor (GlnV44 / supE44) / recombination deficient (recA1) |
|  | P90c**(a)** | ara ∆(lac-pro) thi | Non-Amber suppressor |
|  | MG1655 | K-12 F– λ– ilvG– rfb-50 rph-1 | Non-Amber suppressor |
| **Bacteriophage** |  |  |  |
|  | ʎgt11  (commercially available) | ʎlac5 cIts857 Sam100 LacZ | Blue plaques (IPTG/XGal) |
|  | ʎgt11/Q-am44 | ʎlac5 cIts857 Sam100 Qam44 LacZ | Blue plaques (IPTG/XGal)  Conditionally replicating in supE44 host |
|  | ʎgt11/Q-am44/HDL\* | ʎlac5 cIts857 S-(Gly100amber) Q-(Gln44amber) ∆LacZ | Colorless plaques (IPTG/XGal)  Conditionally replicating in supE44 host |
| **E.coli + Bacteriophage** |  |  | **Relevant phenotype** |
|  | ʎgt11 in DH5α/JM103 |  | Lysis ~2h after induction at 42º (because of Sam100 mutation) **(a)** |
|  | ʎgt11 in MG1655/P90c |  | Lysis ~2 h after induction (42º) **(a)** |
|  | ʎgt11/Q-am44 in DH5α/JM103 |  | Lysis ~2h after induction (42º) **(a)** |
|  | ʎgt11/Q-am44 in MG1655/P90c |  | No lysis after induction (42º) **(a)** |
|  | ʎgt11/Q-am44/HDT in DH5α/JM103 |  | Lysis ~2h after induction (42º) or according to HDL synthetic device |
|  | ʎgt11/Q-am44/HDT in MG1655/P90c |  | Lysis according exclusively to HDL synthetic device |
| \*HDT = Hijack, Detect and Terminate synthetic device  (a) Lin, C. S., et al. 1998 | | | |

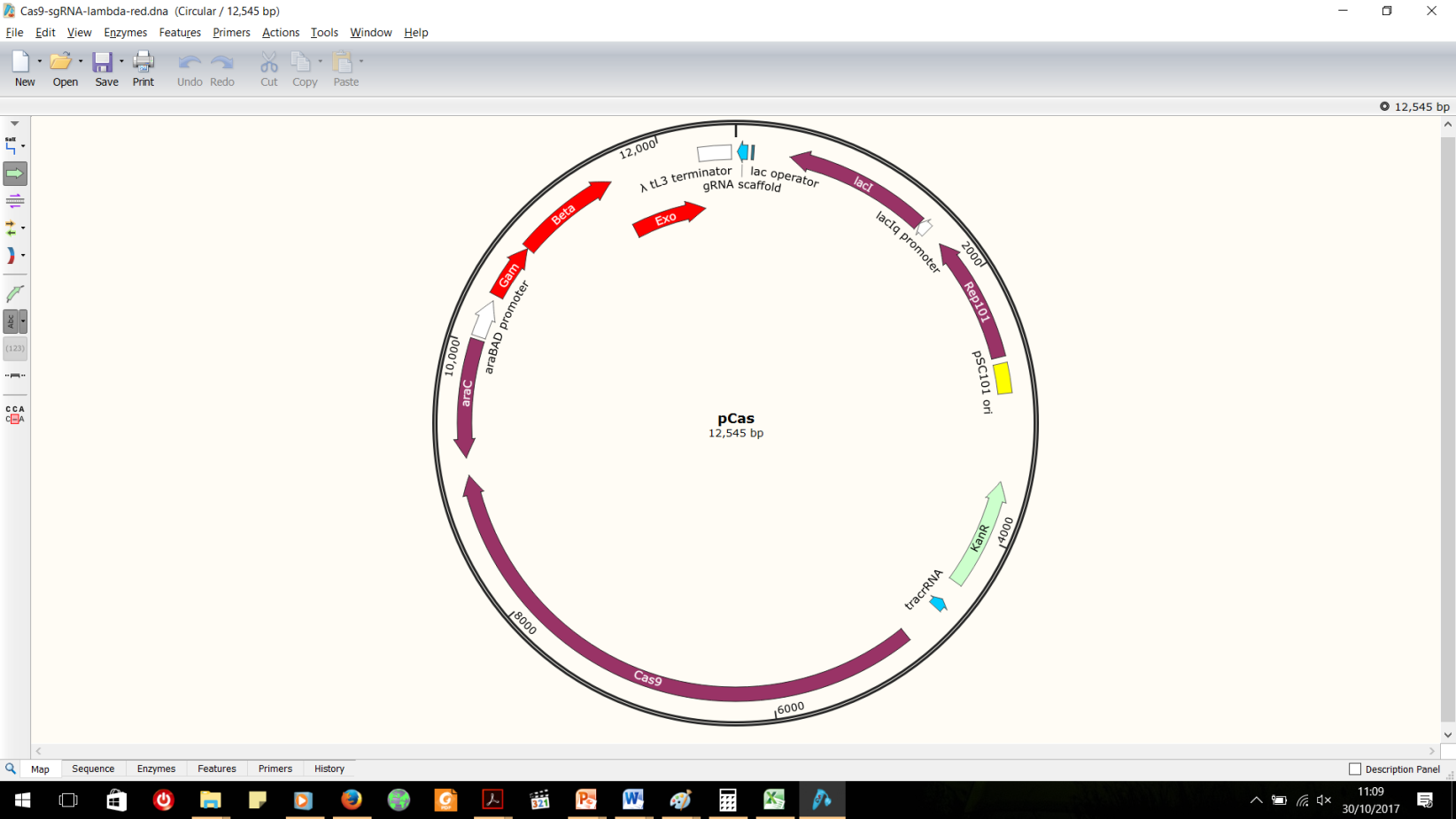
As shown in the Table 1, the most suitable strain for its use in phage engineering is whether DH5α or JM103. The main reason is that these are both amber-supressor strains that are able to complement the mutation of the Q gene while being engineered by CRISPR/Cas9. The **Figure 1** schematizes the procedure for converting ʎgt11 into ʎgt11/Q-am44.



**Figure 1.** Process for the generating a conditionally replicating ʎgt11Q-am44 bacteriophage in a strain of E. coli enabled to suppress the amber mutation (i.e. DH5alpha is not suitable here because contains racA- unfavouring homologous recombination). This E. coli host is co-expressing the CRISPR/Cas9-ʎred plasmid together with a donor plasmid with 800 bp of length of homologous arms (homo, (4) in the Figure). (1) The ʎgt11Q-am44 phage infects the host strain and introduces its genetic material. (2) The CRISPR/Cas9 expression system expresses the Cas9 nuclease and the sgRNA that recognizes the sequence of the wild-type gene Q encoding Gln44. (3) The Cas9 cleaves the sequence upstream the ‘NGG recognition site driven by the sgRNA. (4) The donor plasmid allows for homology directed repair because of the homologous insert contains the fragment substituting the Amber mutation for the Gln44 codon, this can be aided by the Lambda-red recombination system present in the Cas9 plasmid. (5) The ʎgt11Q-am44 is ready for package when it is hosted in a supE44 mutant (i.e. DH5α) providing the Gln aminoacid for an Amber codon and thus allowing the normal packaging of the virus. (6) The ʎgt11Q-am44DNA is packed into the viral capsids. (7) Cell lysis and release of ʎgt11Q-am44 viruses. (8) Plaque formation if ʎgt11Q-am44 has been modified in a supE44-Gln host. (8) Amplification of recovered plaques. (9) Confirmation of the Gln44->Amber mutation by sequencing with appropriate primers.

It is essential to do the procedure for the integration of the amber mutation within a strain of *E. coli* supressing the Amber mutation that will restore the protein to normal levels. Because of that, the strain of Escherichia coli DH5α (or JM103) is a suitable strain because of its supE44 genotype which allows to recover a wild-type Q phenotype from the Qam44 mutant by suppressing the amber mutation (Eggertsson, G., & Söll, D. 1988; Forns, X., Bukh, et al. 1997 ; Singaravelan, B., et al. 2010). It is important to notice that the Amber codon in DH5alpha will be recovered by a Glutamine because of the supE44 genotype and will restore completely the virus functionality.

It is also important to notice that, since the DH5α strain is a *recA* deficient phenotype (i.e. having reduced homologous recombination character), it is necessary to include into the CRISPR plasmid the Red lambda recombination. Fortunately, Jiang, Y., et al. 2015 has constructed a CRISPR/Cas9 plasmid that contains the inducible lambda red recombination system in order to maximise HDR (homology directed repair) after DBS (double strand break) in *E coli* (This plasmid is available in AddGene ID: 62225 <https://www.addgene.org/browse/sequence/155998/> Add hyperlink). See **Figure 2** for a scheme of the CRISPR/Cas9/ʎ-red plasmid visualised in SnapGene®.



**Figure 2.** CRISPR/Cas9 plasmid for the genome editing of the ʎgt11 phage. The plasmid contains the expression system for the Cas9 nuclease and the sgRNA together with a module for the expression of the ʎ-red recombination system (arrows in red). Addgene plasmid # 62225 Add hyperlink

For the Amber mutant strategy to work out, it is necessary to direct the CRISPR/Cas9 system to a region of the gene containing both an aminoacid encoding for Glutamine (i.e. Gln44) and a nearby (<10bp) PAM ‘NGG sequence for Cas9 recognition and efficient HDR. Through the aid of the CRIPSY-web tool (Blin, K., et al. 2016), it has been possible to find a single sgRNA that not only satisfies these two conditions but also has no possible mismatching sequences throughout the bacteriophage genome. This ensures a minimized off-target effect. The CRIPSY-web tool calculates the number of off-target hits containing 0, 1, 2, or 3 bp mismatches on the whole bacteriophage genome for the 13 bp sequence upstream to the PAM sequence by considering all possible combinations within the genome.

The following Table 2 contains the sgRNA guide that can be used to cleave the Q gene in the site mentioned in the paragraph above.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 2.** Spacer sequence for sgRNA guides to be used to generate the Q amber mutant (Qam44) in the ʎgt11 bacteriophage through DSBs with Cas9. The spacers do not contain the adenine methylation site 5’ GATC 3’. All primers have no mismatches (off-target minimized by CRISPY-web algorithm, see Reference in text) | | | |
| #Spacer | Gene | Sequence (5’-3’) 20bp | CG(%) |
| (+) sgRNA for Q | Q | (+)CGTGGTGTGGCAAAGCTTGA | 55 |

After cleavage, in regard to the HDR, although the delivery of linearized ssDNA/dsDNA together with the CRISPR/Cas9 system has been proved to be a successful alternative to modify *Escherichia coli* genomes, as it is much simpler, faster and effective than co-expressing a donor plasmid with long homology arms, this approach is here discarded and instead, a donor plasmid with long homology regions is considered because of the difficulties that would implicate for an infecting phage to electroporated ssDNA/dsDNA together with the CRISPR/Cas9 plasmid *a posteriori*. Therefore, for the modification of the ʎgt11 genome, it is here proposed to provide the phage with a bacterial host that is already prepared to receive the incoming ʎgt11and transform it into the ʎgt11/Q-am44 by means of a donor plasmid that contains 800bp homology arms, although it is clear that the former approach may be feasible and therefore worth a try because of its decrease in the cost. These homology arms, containing the mutation for Qam44, have been designed based on the wild-type version of the bacteriophage ʎ genome (since the complete sequence of ʎgt11 is not available) upon the assumption that the region surrounding gene Q is essential and equivalent between ʎ and ʎgt11. The following **Table 3** summarizes the sequences of the homology arms together with the primers that can be used for its amplification in the ʎ-wild type bacteriophage or in ʎgt11. This homology arms include the Gln44->Amber mutation. The primers that amplify the homology arms include also the biobrick prefix and suffix for enabling its storage in the Biobrick-like backbone pSB1C3 that is suitable because of its high copy number and easiness for which to replace inserts through the biobrick assembly system. The characteristics of pSB1C3 and its relationship with the CRISPR/Cas9 plasmid are summarized in the **Table 4**.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 3.** Homology arms (H. Arm) of 800 bp each (FWD and REV), flanking an Amber insertion (TAG), used to generate the Q- amber mutant (Gln44-Amber) of ʎgt11 conditionally replicating phage. All primers have been designed to contain a difference in melting poing (Tmº) lower than 2º. Primers and melting points have been calculated from the SnapGene® software to ensure no mismatches/off-target throughout the genomic sequence. | | | |
| #H. Arm | Gene | Homology sequence (5’-3’) 800 bp  Primers (5’-3’) with Tm (melting Tº) | 5’NGG3’ distance (bp) |
| 1A | Q  (upstream /  H-left) | 5’gacctgcttatctcggtgggcgatttggttgatcgtggtgcagagaacgttgaatgcctggaattaatcacattcccctggttcagagctgtacgtggaaaccatgagcaaatgatgattgatggcttatcagagcgtggaaacgttaatcactggctgcttaatggcggtggctggttctttaatctcgattacgacaaagaaattctggctaaagctcttgcccataaagcagatgaacttccgttaatcatcgaactggtgagcaaagataaaaaatatgttatctgccacgccgattatccctttgacgaatacgagtttggaaagccagttgatcatcagcaggtaatctggaaccgcgaacgaatcagcaactcacaaaacgggatcgtgaaagaaatcaaaggcgcggacacgttcatctttggtcatacgccagcagtgaaaccactcaagtttgccaaccaaatgtatatcgataccggcgcagtgttctgcggaaacctaacattgattcaggtacagggagaaggcgcatgagactcgaaagcgtagctaaatttcattcgccaaaaagcccgatgatgagcgactcaccacgggccacggcttctgactctctttccggtactgatgtgatggctgctatggggatggcgcaatcacaagccggattcggtatggctgcattctgcggtaagcacgaactcagccagaacgacaaacaaaaggctatcaactatctgatgcaatttgcacacaaggtatcggggaaataccgtggtgtggcaaagcttTAG3’  **FWD Primer 5’3’ – with biobrick-prefix**  GAATTCGCGGCCGCTTCTAGAGgtgaaagagatgcgctattacgaaaaaattgatggcagc (Tm-62º)  **REV Primer 5’3’ – with TAG (Q- Gln44->Amber mutant)**  CTAtgattgcgccatccccatagca (Tm-62º) | -2 |
| 1A | Q  (downstream / H-right) | **5’TAG**gccggattcggtatggctgcattctgcggtaagcacgaactcagccagaacgacaaacaaaaggctatcaactatctgatgcaatttgcacacaaggtatcggggaaataccgtggtgtggcaaagcttgaaggaaatactaaggcaaaggtactgcaagtgctcgcaacattcgcttatgcggattattgccgtagtgccgcgacgccgggggcaagatgcagagattgccatggtacaggccgtgcggttgatattgccaaaacagagctgtgggggagagttgtcgagaaagagtgcggaagatgcaaaggcgtcggctattcaaggatgccagcaagcgcagcatatcgcgctgtgacgatgctaatcccaaaccttacccaacccacctggtcacgcactgttaagccgctgtatgacgctctggtggtgcaatgccacaaagaagagtcaatcgcagacaacattttgaatgcggtcacacgttagcagcatgattgccacggatggcaacatattaacggcatgatattgacttattgaataaaattgggtaaatttgactcaacgatgggttaattcgctcgttgtggtagtgagatgaaaagaggcggcgcttactaccgattccgcctagttggtcacttcgacgtatcgtctggaactccaaccatcgcaggcagagaggtctgcaaaatgcaatcccgaaacagttcgcaggtaatagttagagcctgcataacggtttcgggattttttatatctgcacaacaggtaagagcattgagtcgataatc 3’  **FWD Primer 5’3’ – with TAG (Q- Gln44->Amber mutant)**  TAGgccggattcggtatggctgc (Tm63º)  **REV Primer 5’3’ – with biobrick-sufix**  CTGCAGCGGCCGCTACTAGTAgattatcgactcaatgctcttacctgttgtgc (Tm62º) | -2 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 4.** Compatible plasmids co-expressed in E coli DH5alpha for genome editing. | | | | |
| **Plasmid** | Ori | Resistance | Copy number | Length (bp) |
| CRISPR/Cas9/lambda-red (Addgene #62225) | pSC101 | Kanamycin | Low | 12545 |
| Biobrick (pSB1C3) + donor | ColE1/pMB1/pBR322/pUC | Chloramphenicol | High | 2070 + 1603 |

Once the biobrick (i.e. pSB1C3) incorporating the donor sequence (i.e. homology arms containing gene Qam44 flanked by the biobrick suffix and prefix), it is possible to co-transform this plasmid together with the CRISPR/Cas9/ʎ-red plasmid in order to create the Amber mutant phage, as seen in the Figure 1. The two plasmids are able to co-express in a same *E. coli* host because they contain different replication origins and different Antibiotic markers.

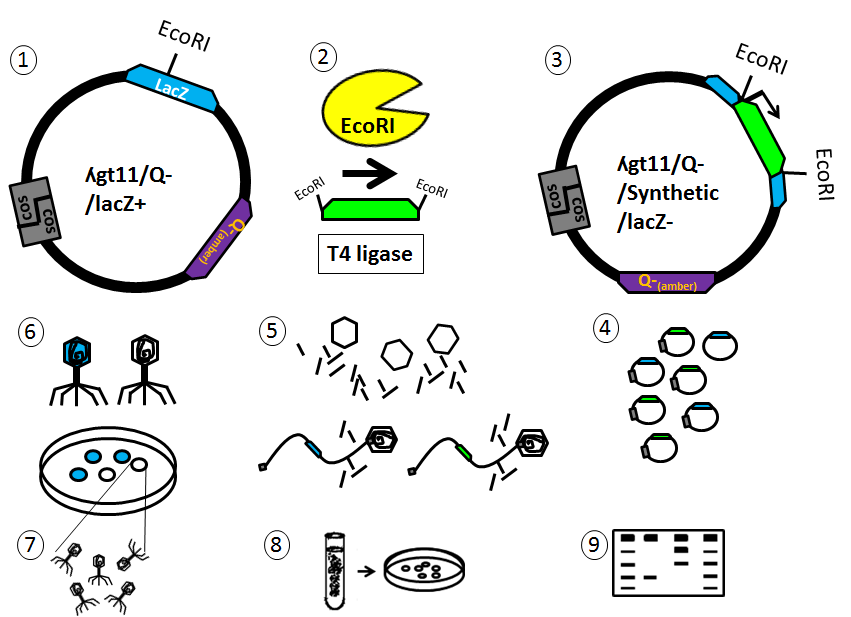
Upon success, the transformed Escherichia coli DH5α (or JM103) containing the CRIPSR/Cas9 and the pSB1C3 plasmid containing the donor homologous sequence will be suitable for infection and this will lead to the generation of ʎgt11Qam44 mutants. The CRIPSR/Cas9 will be actively performing the cribbage for the wild-type versions of the phage (see Figure 1) so that only mutants will be expected as an output.

After that, the new ʎgt11Q-am44 phage can be sequenced with primers flanking the homology arms and a conditionally replicating ʎgt11Q-am44 bacteriophage will be confirmed. Next step will be to test it in a suitable non-amber-suppressor strain of *E. coli*, such as MG1655 or P90c (see Table 1).

Finally, this new phage will have these four main characteristics:

1. Termosensible (*cI857* mutant). Lytic phage is activated at 42º. Lysogens stay at 32º.
2. Conditionally replicating in DH5α or JM103 (contain the supE44 genotype)
3. Harbour a space of up to 9625 bp in a unique EcoRI site in the LacZ gene
4. If it does not contain an insertion in EcoRI site, the phage is LacZ positive (gives blue plaques in presence of IPTG and XGal). If an insert is present, it is LacZ negative (gives colorless plaques).

With this phage in mind, it will be as easy to implement synthetic DNA into ʎgt11Q-am44 as it is to modify the TM4 vector (phAE159). It is only necessary that the synthetic DNA is flanked by EcoRI sites. The **Figure 4** below schematizes the procedure to implement synthetic DNA into the ʎgt11Q-am44.



**Figure 4.** Procedure for the incorporation of a synthetic device into **ʎgt11/Q-**. (1) Isolation of **ʎgt11/Q-/lacZ+** with unique EcoRI site. (2) Digestion of **ʎgt11/Q-/lacZ+** with EcoRI together with the synthetic insert (green). (3) ligation of the synthetic insert to give rise to **ʎgt11/Q-/lacZ-/Synthetic**. (4) Preparation of the resulting plasmids for packaging in vitro. (5) Packaging in vitro mediated by *cos* sites (6) Infection and plaque assay in an “Amber supressor and LacZ- E. coli stain (i.e. DH5alpha)” with the resulting packed phages. (7) Plaque isolation of the colorless plaques harbouring the synthetic DNA (i.e. LacZ-). (8) Amplification of the synthetic phage. (9) Confirmation by sequencing.

In order to perform the procedure schematized in the Figure 4, it is very useful to check the comprehensive protocols available for the integration of inserts into ʎgt11 (see References: Lambda gt11/EcoR I/CIAP-Treated Vector Kit – Instruction Manual) or (Jendrisak, J., et al. 1987) which provides as well a very detailed way in which to implement synthetic DNA into ʎgt11-like vectors.

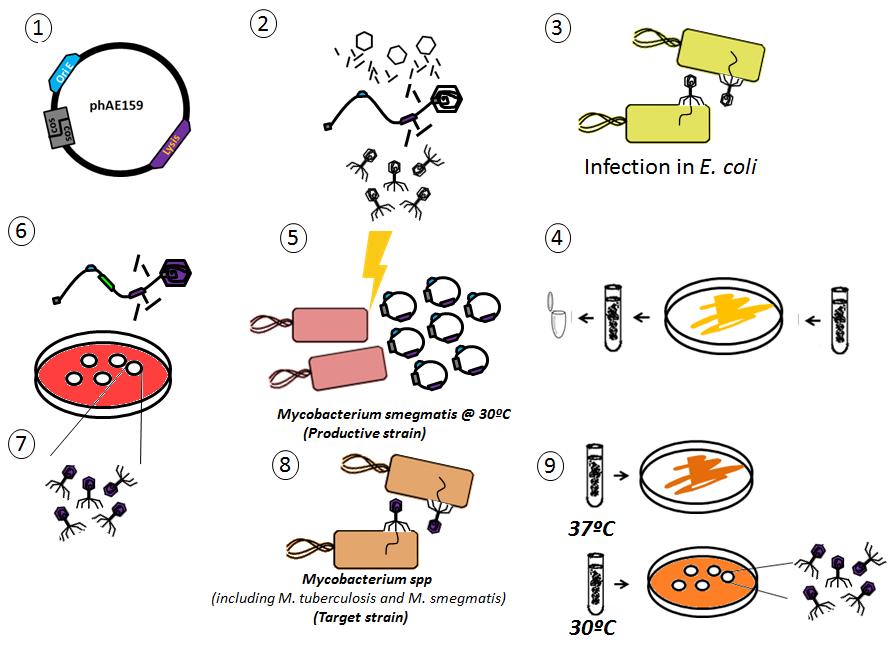
This can be used to definitely include the Hijack, Detect and Terminate system in order to test its functionality in *Escherichia coli* *MG1655* or *P90c*, since they are not amber supressing mutants and the gene Q will be deactivated, giving rise to a new version of the phage that will behave according to the synthetic DNA, as a way to control the cell cycle, to detect it with colour dyes or fluorescence, and finally lyse according to the different terminators designed.

**REFERENCES**

* Blin, K., Pedersen, L. E., Weber, T., & Lee, S. Y. (2016). CRISPy-web: an online resource to design sgRNAs for CRISPR applications. *Synthetic and Systems Biotechnology*, *1*(2), 118-121.
* Chauthaiwale, V. M., Therwath, A., & Deshpande, V. V. (1992). Bacteriophage lambda as a cloning vector. *Microbiological reviews*, *56*(4), 577-591.
* Eggertsson, G., & Söll, D. (1988). Transfer ribonucleic acid-mediated suppression of termination codons in Escherichia coli. *Microbiological reviews*, *52*(3), 354.
* Garrett, J., Fusselman, R., Hise, J., Chiou, L., Smith-Grillo, D., Schulz, J., & Young, R. (1981). Cell lysis by induction of cloned lambda lysis genes. *Molecular and General Genetics MGG*, *182*(2), 326-331.
* Forns, X., Bukh, J., Purcell, R. H., & Emerson, S. U. (1997). How Escherichia coli can bias the results of molecular cloning: preferential selection of defective genomes of hepatitis C virus during the cloning procedure. *Proceedings of the National Academy of Sciences*, *94*(25), 13909-13914.
* Gelvin, S. (2012). *Plant molecular biology manual*. Springer Science & Business Media.
* Jacobs Jr, W. R. (2014). Gene transfer in Mycobacterium tuberculosis: shuttle phasmids to enlightenment. *Microbiology spectrum*, *2*(2).
* Jendrisak, J., Young, R. A., & Engel, J. D. (1987). [40] Cloning cDNA into λgt10 and λgt11. *Methods in enzymology*, *152*, 359-371.
* Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., & Yang, S. (2015). Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system. *Applied and environmental microbiology*, *81*(7), 2506-2514.
* Knipe, D. M., Howley, P. M., Griffin, D. E., Lamb, R. A., Martin, M. A., Roizman, B., & Straus, S. E. (2007). Fields virology, vol. 1. 5th Edition. *Philadelphia (EUA): Lippincott Williams & Wilkins*, a Wolters Kluwer Business
* Lambda gt11/EcoR I/CIAP-Treated Vector Kit – Instruction Manual - [http://hpst.cz/sites/default/files/attachments/g11.pdf Retreived 20/09/17](http://hpst.cz/sites/default/files/attachments/g11.pdf%20Retreived%2020/09/17)
* Lin, C. S., Chen, B. Y., Park, T. H., & Lim, H. C. (1998). Characterization of bacteriophage λ Q− mutant for stable and efficient production of recombinant protein in Escherichia coli system. *Biotechnology and bioengineering*, *57*(5), 529-535.
* Moran, L. S., Maina, C. V., Poole, C. B., & Slatko, B. E. (1990). Nucleotide sequence of the phage λgt11 SacI-KpnI lacZ region. *Gene*, *93*(1), 163-164.
* Singaravelan, B., Roshini, B. R., & Munavar, M. H. (2010). Evidence that the supE44 mutation of Escherichia coli is an amber suppressor allele of glnX and that it also suppresses ochre and opal nonsense mutations. *Journal of bacteriology*, *192*(22), 6039-6044.

**ENGINEERING phAE159 TO HIJACK, DETECT AND TERMINATE M. TUBERCULOSIS**

The strategy proposed here for the design of a mycobacteriophage able to infect *Mycobacterium tuberculosis* so that it can express the Hijack, Detect and Terminate system relies on the work started by Jacobs, W.R. Jr, et al. 1987 in which he constructs shuttle plasmids that replicate as large plasmids (cosmids) in *E. coli* and as mycobacteriophages in *Mycobacterium spp*. These shuttle phasmids are constructed by ligation of the mycobacteriophage DNA (here the TM4 mycobacteriophage that infects *Mycobacterium spp*, including *M. tuberculosis*) to a cosmid vector containing an origin of replication (OriE) for *E. coli* and *cos* sites required for packaging into phage lambda particles in vitro. The **Figure 5**, schematizes the way in which phAE159 shuttle vector behaves in both *E. coli* and *Mycobacterium spp* under different conditions that are characteristic. The phAE159 can be used to infect E. coli, where it can be amplified as a cosmid, and this DNA can be isolated and then used to transform the strain of Mycobacteria that growth faster, such as *M. smegmatis*. This shuttle procedure has been demonstrated to be useful for recovering plaques with TM4-like bacteriophages that retain the cosmid part of the vector (Jacobs, W.R. Jr, et al. 1989). This relies on the fact that phAE159 has an appropriated genomic/cosmid size for packaging in both ʎ- and TM4-capsids. The packaging constraints for phage lambda are ~53000 bp and those for TM4 are still not determined. The plasmid sequence of phAE159 can be taken from the supplementary material of: Jain, P., et al. 2014.

****

**Figure 1.** The phAE159 shuttle vector is an essential mycobacteriophage which a plasmid/cosmid region containing *OriE* for *E. coli* replication, and *cos* sites for enabling package in ʎ capsids. (1) The phAE159 plasmid can (2) be packed in vitro in ʎ-packaging kits for infection in *E. coli*. (3) ʎ-packed phAE159 is used as a vector to deliver the plasmid (cosmid) into *E. coli*. (4) The result of the infection can be plated to test survival from infection and it can be further amplified for isolation and storage of large quantities of the cosmid. (5) Transduction of Mycobacterium smegmatis (grows faster and is used as phage production strain) with the amplified cosmids by electroporation. (6) The cosmid, as a TM4-derived Mycobacteriophage, will assemble into its capsids in Mycobacterium spp and will form plaques at 30º (not at 37º since phAE159 is termosensible). (7) High titer TM4-derived phAE159 mycobacteriophage can be recovered from plaques. (8) The recovered phAE159 mycobacteriophages can be used to easily infect any species of Mycobacteria (such as the slow growing *M. tuberculosis*). (9) At 37º the termosensible phAE159 TM4-derived mycobacteriophages will not produce plaques, whereas at 30ºC mycobacteriophages will be released in form of plaques.

As shown in the Figure 1 above, the phAE159 is already a termosensible conditionally replicating TM4-derived mycobacteriophage that does not complete the viral cycle when it is at 37ºC. If a synthetic device is included in phAE159, and this device does contain promoters that are not regulated by the mycobacteriophage (such as the pHyb promoter) but that are able to be expressed strongly in Mycobacteria, the expression of this synthetic device is expected not to be inactivated, regardless of the induction of the virus or not. Nevertheless, the level of expression will be strongly affected by whether the virus will be induced (at 30º) or not (37º), see **Figure 5**. Induced viruses will replicate voraciously and therefore the copy number of the synthetic device will be higher than for the non-induced state, therefore those induced will present higher levels of expression. The non-induced virus will not produce high levels of expression since the virus will not produce copies of itself once within its host. The only inconvenient of an induced TM4-derived mycobacteriophage is that the lysis of the cell will not enable for sustained expression of the synthetic genes incorporated with phAE159 and therefore a solution for this issue is inhibiting the lysis of the host without compromising the increase in copy number by induction of the mycobacteriophage and still conserving the possibility to produce progeny viruses.

The solution for avoiding lysis, without compromising the possibility to produce mycobacteriophages for its use as a detection system, is by generating a mycobacteriophage that is as well conditionally replicating in regard to the Lysis genes (LysA, LysB and the here predicted Holin gene). Both LysA and LysB genes are well documented in the literature but the Holin gene has not been annotated in the databases, although it can be predicted *blasting* its the protein sequence downstream of LysA and LysB. Marinelli, L.J., et al. 2008 demonstrates that it is possible to generate a strain of Mycobacteriophage Giles where essential lysin A (lysA) gene is deleted and it can be recovered by complementation strains expressing the Lysis genes in *trans*. Although the paper claims LysA is essential and it even demonstrates it, it is advisable to knock down LysA altogether with LysB and the Holin gene as well since a controversial study of *Payne, Kimberly., et al 2009* shows the host dynamics whenLysA or LysB mycobacteriophage mutants are employed and appears that it is still possible produce progeny viruses although it is clear that lysis is severely impaired. Therefore, the deletion of all of the genes involved in the Lysis (LysA, LysB and Holin) in the TM4-derived phAE159 is proposed here a secure approach for the prolonged expression of the synthetic device, ensuring that it conserves the possibility to produce Mycobacteriophages by conditionally replicating them in suitable hosts expressing the lysis genes. As an example, Lin, C.S., et al 1998 shows that blocking lysis in temperate induced ʎ-bacteriophage enables for much higher gene expression levels than for the temperate non-induced version (i.e. lysogenic). *Tanji, Y., et al. 2004* demonstrates how a lysozyme-inactivated T4 bacteriophage can be used as a detection system.

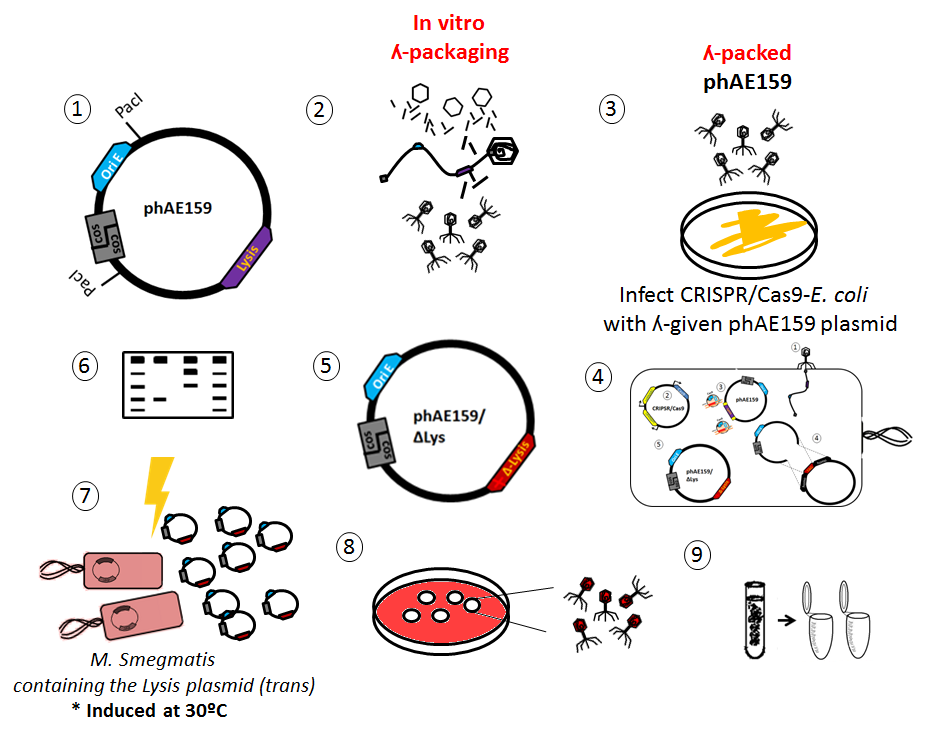
For this particular Hijack, Detect and Terminate synthetic device, there is another constraint in phAE159: the inclusion of a synthetic construct as large as ~9000 bp. Therefore, to determine whether a synthetic construct of ~9000 bp can be hosted in TM4-derived phAE159 bacteriophage it is important to have the following considerations:

* The wild-type TM4 bacteriophage has genome length of 52797 bp, therefore, since there is no available literature stating maximum package limits in TM4, the maximum packaging capacity for phAE159 is 52797 bp. (Hyperlink to genome file: TM4 from U pittspburgh (featured).dna )
* phAE159 has a genomic length of 50725 bp, (Hyperlink to genome file: phAE159\_featured\_with\_sgRNA.dna) leaving space for 2072 bp of synthetic DNA. Furthermore, it is possible to remove the cosmid vector that is flanked by unique PacI restriction sites (see **Figure 5**) containing the *OriE* and *cos* sites, releasing an additional 3753 bp. Therefore the total amount available in a normal phAE159 is 5825 bp. This amount is still far from the 9000 bp
* The removal of the Lysis cassette from phAE159 gives an additional available space of 3261 bp, making possible a total available space of 9086 bp.

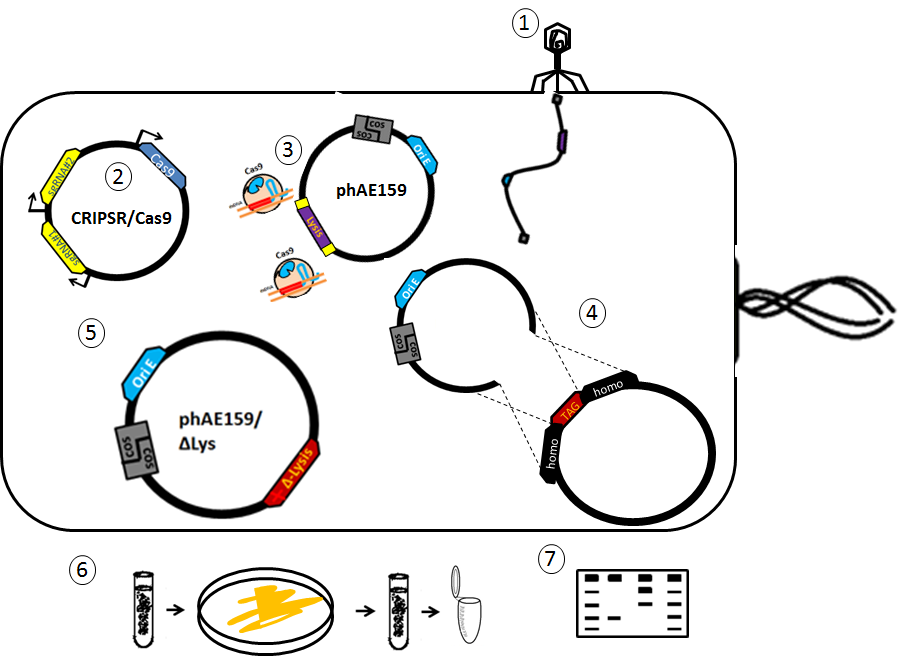
Therefore, the generation of a Lysis-conditionally replicative TM4-derived phAE159 not only enables to use phAE159 as a Hijack, Detect and Terminate system by hampering on the lysis of the host, but also is the most secure alternative to host the synthetic construct without compromising package in TM4 capsids.

**1st. Generating a ∆LysA, ∆LysB, ∆Holin (∆Lys) conditionally replicating mycobacteriophage from phAE159 in Escherichia coli**

The CRISPR/Cas9 gene editing system has not yet been optimized for its direct use in *Mycobacterium spp*. Because of that, since phAE159 disposes of origins of replication for both *Escherichia coli* (OriE) and *Mycobacterium spp* (OriM, not shown) it is possible to use this shuttle property to edit phAE159 in *E. coli* as a plasmid prior to its propagation in *Mycobacterium spp* as a mycobacteriophage. The **Figure 2** represents a scheme of the procedure.



**Figure 2.** Gene editing of the phAE159 to generate the ∆LysA, ∆LysB, ∆Holin conditionally replicating variant is done in *Escherichia coli*, whereas the production of mycobacteriophage is done upon expression in *Mycobacterium spp*. (1) The phAE159 containing the cos sites for ʎ-packaging and OriE for replication in *E. coli* is isolated and purified from its source. (2) The purified phAE159 is packed in vitro in ʎ bacteriophage (white). (3) ʎ-packed phAE159 infects *Escherichia coli* containing both the CRIPSR/Cas9 plasmid targeting the Lysis cassette (LysA, LysB, Holin genes) in phAE159, and the donor plasmid containing the TAG (amber sequence) flanked by large homology arms surrounding the former Lysis cassette. (4) ∆Lys-phAE159 is generated in *Escherichia coli* following the procedure of the Figure 3. (5) ∆Lys-phAE159 obtained by homologous recombination with the donor plasmid in *E. coli*. (6) The resulting ∆Lys-phAE159 plasmid can be recovered using a conventional cosmid/plasmid extraction protocol. (7) The purified ∆Lys-phAE159 plasmids are transduced into *Mycobacterium smegmatis* already containing a complementing inducible Lysis cassette (notice *M smegmatis* already containing a plasmid), by electroporation at 30ºC to induce the viral genes (phAE159 is already a conditionally replicating termosensible mycobacteriophage) and enable the viral cycle in a plaque assay. (8) Isolation of TM4-packed ∆Lys-phAE159 from high titer plaques. (9) Storage of the contidionally replicating TM4-packed ∆Lys-phAE159 for its use as a Hijack, Detect and Terminate system in Mycobacteria.



**Figure 3.** CRIPSR/Cas9 is used to generate ∆Lys-phAE159 in *Escherichia coli*. (1) The ʎ-packed phAE159 infects *E. coli* and allows for the stable replication of phAE159 within this host. (2) A constitutively expressed plasmid, already present upon infection, with CRISPR/Cas9 system targeting the Lysis casette in phAE159 (here in violet). (3) The phAE159 plasmid is cleaved by Cas9 directed by the sgRNA guides at specific sites flanking the Lysis cassette (cassette shown in violet, flanked by the sites in yellow). (4) A donnor plasmid containing a TAG (amber sequence) flanked by large homology arms allows for the incorporation of the deleting sequence by means of homologous recombination of the cleaved site in phAE159. (5) phAE159-∆Lys conditionally replicating mycobacteriophage is hosted as a replicative plasmid in *E. coli* because it still contains the *OriE*. Wild-type versions of the phAE159 are not expected since these are cribbed in by the constitutive expression of CRIPSR/Cas9. (6) The infected *E. coli* cells can be plated as a survival test, amplified, and further processed by Miniprep® to isolate the plasmid DNA. (7) The ∆Lys-phAE159 plasmid can be sorted by size exclusion in an agarose gel.

For the deletion of the Lysis cassette spacer sequence for the CRISPR/Cas9 system have been designed by using the CRIPSY-web tool (Blin, K., et al. 2016) in order to make sure there are no off-target effects. The following **Table 1** contains the sgRNA guides that can be used to cleave the Lysis cassette (LysA, LysB, Holin). Upon cleavage, a insertion of a TAG (amber mutation) will be substituted by homologous recombination by using a donor plasmid whose homology arms are stated in the **Table 2**.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 1.** Spacer sequence for sgRNA guides to be used to generate the ∆LyA,∆LysB,∆Holin mutant in the phAE159 mycobacteriophage through DSBs with Cas9. The spacers do not contain the adenine methylation site 5’ GATC 3’. All primers have no mismatches (off-target minimized by CRISPY-web algorithm, see Reference in text) | | | |
| #Spacer | Gene | Sequence (5’-3’) 20bp | CG(%) |
| (+) sgRNA #1 for Lysis cassette | LysA, LysB, Holin | (+)GAGTTTCACCCGGTTCCTGC | 60 |
| (+) sgRNA #1 for Lysis cassette | LysA, LysB, Holin | (-)GGCATGCTGGCTCGTGCTGA | 65 |

These fragments can be amplified directly from phAE159 or the wild-type TM4 genome by using the corresponding primers also stated in the Table 2 below.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 2.** Homology arms (H. Arm) of 800 bp each (FWD and REV), flanking an Amber insertion (TAG), used to generate the ∆LyA,∆LysB,∆Holin mutant in the phAE159 conditionally replicating mycobacteriophage. These fragments can be amplified from phAE159 or the wild-type TM4 genome. All primers have been designed to contain a difference in melting poing (Tmº) lower than 2º. Primers and melting points have been calculated from the SnapGene® software to ensure no mismatches/off-target throughout the genomic sequence. These constructs generate a deletion of 3261 bp | | | |
| #H. Arm | Gene | Homology sequence (5’-3’) 800 bp  Primers (5’-3’) with Tm (melting Tº) | 5’NGG3’ distance (bp) |
| 1A | Lysis cassette  (upstream /  H-left) | 5’ggtggcgcaggcagataacccgcacgcccgcgggttcgccatcgtctactcgcccaaggtcacttggcggccctggccgcccgccgacgactgagcgcatgctcgctggccctgataacggccgagcgggtttcgtagcacgtgcgcctcgcccgtagctaaccgccggcaaaaactcgagcaccccctctcacctgcggatacctcgcgaaacggcgcgaggtcggcaaacactggataggagcaccgtggcagcaacggacgcattcaagctcgcgatcgccaacgcgatcggcgcgcaaggcgcactgatcagcctgcactcggctgaccccggcaagaccgacgcgacggccaacgcgaccgaaattagcggcgccggatacacgcgcaagctgaccgcgtggggcgccccggtcatcgtgtcgggcggcgccgacgacggcaaggcccgcatcaccggctcgacgcagcagttcaacgtgcccggcggcgtgccgatcacgcactacgccgtgcgcaaggccgacagcacattcctgtacggcaagccgctggcgcccggcgcgaccctcaccggcaacggtgtcatcgacgtcacgccgacacatacctacgacctgacctagctcgaaatggtcggcgtcgagggcattttcgcagcattgtctgcggctgtggtgctcggcgccctcgggcactggctctatgacgtgctggcgcaccggcgctacgacaacgacgagggatacgacacatgagtttcacccggttTAG3’  (774bp)  **FWD Primer 5’3’ – with biobrick-prefix**  GAATTCGCGGCCGCTTCTAGAGggtggcgcaggcagataaccc(Tm-58º)  **REV Primer 5’3’ – with TAG (LysA,LysB,Holin)**  CTA aaccgggtgaaactcatgtgt (Tm-60º) | -5 |
| 1A | Lysis cassette  (downstream / H-right) | **5’TAG**gctggctcgtgctgacaacgcctcggcggcgaagctggacggcatgcgggggatgccatgacgtaccgctacgtcgagaaccgggtgctgcgtttcgtgcagctcgcgctgctcgtcgacgccgtcgtgcgcggcgcgagctggatcgcgaccccggcgagcggcatacccccagcgatcggcctggccgccgaaggcaccgcagccatgtgggtgtggggcgccgtattcgccgtgtttggcatcttgggcctgctcggcgagctgtggatgcacctcggcgagtctgagcatcgcgcgtggccgtcattcctggcgtacgccgcgctgctgttcctgttcgccgggctggccctgtcggcggtcaataacgtcatcacaacgcacgcaacggacgggttcagcgccccatacactttcgcccttctcgcgttgctgcattgggtgttcgcgaggcggcggaagcatgtcggctgagctgatcgagaagctgccgcagcagtgggtcggcatcgtcgtcctggtgctgttcgtcgtctacgtggccgggcagctcatcgagaaatccgagcgcgtcgcgaagctgctgcccctcggcgtgtggtggcgcgagcgcaaccggcgcaagtctgcggtcgacccggccgagctgacacgcgcggtcgaggcggcccggcatgcctggtcgcgcgaggagaacgccgcactggcggcccttgagagccgcgtcgccgtgatcgccgcgatttcggagcaccaggccctcaacatcaaagagctgcaagactcggtgcgggcgttcacggcgttctctgtgtacgacgcg 3’  (826 bp)  **FWD Primer 5’3’ – with TAG (LysA,LysB,Holin)**  TAGgctggctcgtgctgacaac (Tm61º)  **REV Primer 5’3’ – with biobrick-sufix**  CTGCAGCGGCCGCTACTAGTAcgcgtcgtacacagagaacgc (Tm59º) | -2 |

As stated above, and considering the TAG (3bp) introduced after deletion, a total of 9083 bp is available after deletion of the Lysis cassette, giving enough space to host the two different versions of the Hijack, Detect and Terminate.

Synthetic construct 1 = Hijack, Detection and Termination-LacZa-CcdB = 8105 bp

Synthetic construct 2 = Hijack, Detection and Termination-by-Lysis = 9056 bp

**2nd Conditionally replicating host for the production of engineered synthetic mycobacteriophages**

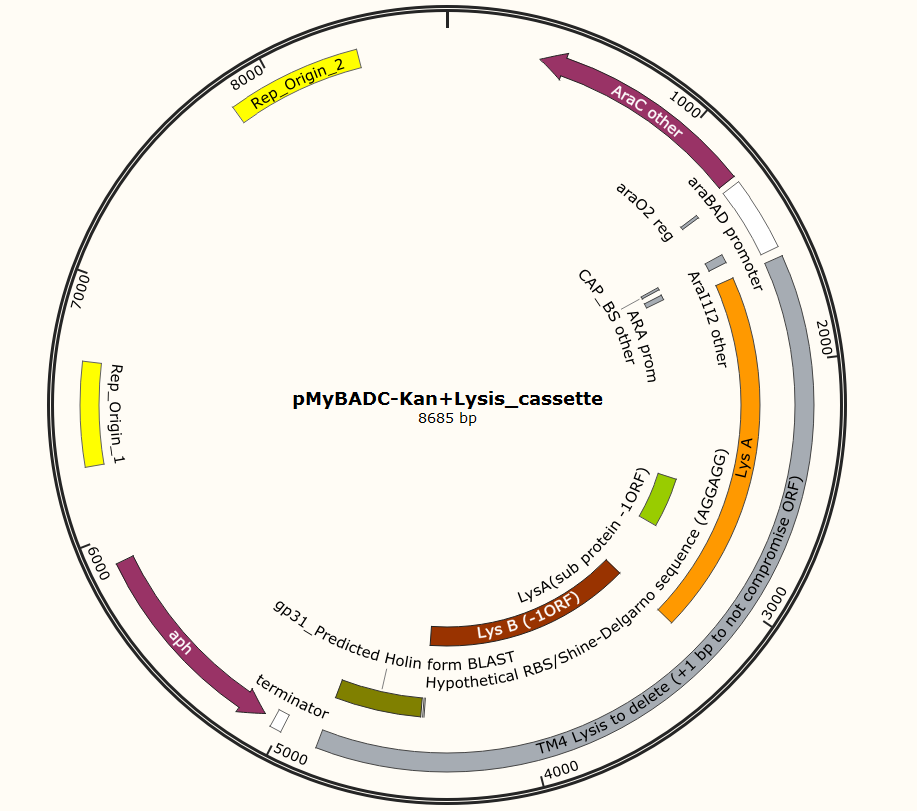
A host will be necessary for the production of mycobacteriophages under suitable conditions not related to the HDT device. Therefore, in order to dispose of a platform for the rapid generation of these engineered mycobacteriophages to be used later as a HDT system in the target *Mycobacterium tuberculosis* it is important to build a non-virulent host strain that will also be able compensate for the deletion in the Lysis cassette by using an expression plasmid. This will enable the mycobacteriophage to be conditionally replicated in the production strain (see Figure 1 and Figure 5). *Mycobacterium smegmatis* can be used for such a purpose since it has a doubling time up to 4h while *Mycobacterium tuberculosis* has a doubling time of 24h.

Available plasmids for mycobacterial expression are the ones from the Matthias Willmans Lab at EMBL Hamburg (<http://www.embl-hamburg.de/research/unit/wilmanns/index.html>). Here it is proposed the use of pMyC-Kan (Addgene plasmid #84692) or pMyBADC-kan (Addgene plasmid #84689) which are plasmids for inducible expression under an acetamidase or AraC promoter, respectively. These plasmids contain appropiate MCS restriction sites for the incorporation of this particular lysis cassette (i.e. available restriction sites).

In order to introduce the lysis cassette into the pMyC-Kan/pMyBADC-kan plasmid, an appropriate approach is to amplify the cassette from the non-engineered phAE159 or the wild-type TM4 genome by using primers flanked by restriction sites present on the MCS site that are not present in the Lysis cassette. In this case, NcoI and EcoRI restriction sites are selected (see pMyC-Kan (addgene link: [www.addgene.org/84692/](http://www.addgene.org/84692/)) and pMyBADC-kan (addgene link: <https://www.addgene.org/84689/> )). The primers that can be utilised for such purpose are stated in the following the **Table 3.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 3.** Primer sequences for the amplification of the Lysis cassette from phAE159. Melting points for FWD and REV primers are close to each other. | | | |
| #Spacer | Gene(s) | Sequence (5’-3’) | Tm(ºC) |
| FWD Primer 5’3’ with NcoI | LysA, LysB, Holin | CCATGGatgagtttcacccggttcctg | 60 |
| REV Primer 5’3’ with EcoRI | LysA, LysB, Holin | GAATTCttgtcagcacgagccagcatg | 60 |

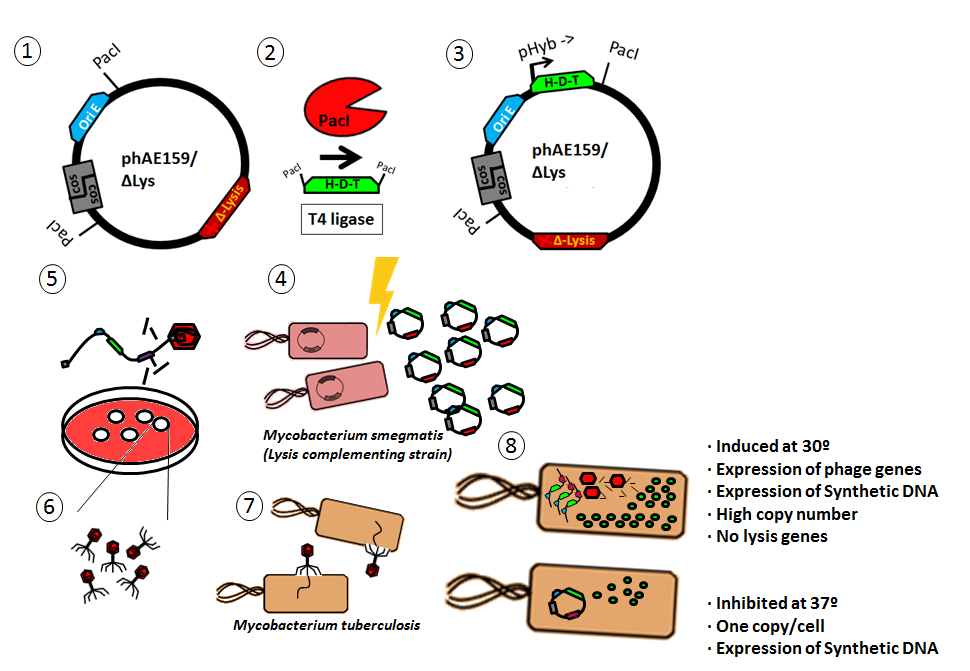
The resulting amplicon by using the primers from Table 3 can be seen here (add hyperlink with SnapGene file: Amplicon\_LysA,LysB,Holin\_NcoI\_EcoRI.dna). The resulting plasmid is the one shown in the **Figure 4**. The sequence can be found here (add hyperlink with the SnapGene file: pMyBADC-Kan+Lysis\_cassette.dna)



**Figure 4.** pMyBADC-kan expression plasmid containing the Lysis cassete (LysA, LysB and Holin genes) from TM4 mycobacteriophage (phAE159). See the sequence here (add hyperlink with filename : pMyBADC-Kan+Lysis\_cassette.dna). The plasmid has been constructed using SnapGene.

In this system, only mycobacteriophages infecting hosts containing the pMyBADC-kan (or pMyC-kan) plasmid with the Lysis cassette and in presence of Arabinose (or acetamidase for pMyC-kan) inductor, will complete normalised viral cycles and will release the viral progeny.

Once a first cycle of the ∆Lys-phAE159-HDT containing the Hijack, Detect and Terminate (HDT) system is produced in the producing strain containing the Lysis cassette the recovered conditionally replicating mycobacteriophages can be finally used as a Hijack, Detect and Terminate system for *Mycobacterium tuberculosis*, see **Figure 5** for more details.

****

**Figure 5.** (1 and 2) The phAE159-HDT (∆LysA, ∆LysB, ∆-holin) will be digested in vitro targeting the unique *PacI* restriction sites. (3) Ligation of the Synthetic construct for Hijack-Detect-Terminate (H-D-T) optionally containing the cos site and oriE for alternative packaging in ʎ and replication in *E. coli* replication. (4) Electroporation of phAE159-HDT into *Mycobacterium smegmatis* harbouring the complementing plasmids (with LysA, LysB and hoin genes). (5) Replication of the Mycobacteriophage and packaging of phAE159-HDT synthetic mycobacteriophages containing the synthetic deivice. (6) Release of new bacteriophages containing the synthetic device. (7) Infection of a non-complementing strain such as the target host *Mycobacterium tuberculosis*. (8) If the prophage is induced at 32º, the viral cycle will be activated and multicopy expression of genes, including capsid formation and packaging will happen, nevertheless because of being in a non-complementing host strain, the lysis of the host is not expected. If the prophage is inhibited by raising the temperature at 37º the bacteriophage will be in a Pseudo-lysogenic state and the expression of the genes controlled by host promoters may be possible in low copy numbers. The inactivation of LysA, LysB and holin genes can be done in Escherichia coli, using CRIPSR/Cas9.

**REFERENCES**

* Jacobs, W. R., Tuckman, M., & Bloom, B. R. (1987). Introduction of foreign DNA into mycobacteria using a shuttle phasmid. *Nature*, *327*(6122), 532-535.
* Jacobs Jr, W. R., Snapper, S. B., Tuckman, M., & Bloom, B. R. (1989). Mycobacteriophage vector systems. *Reviews of infectious diseases*, *11*(Supplement\_2), S404-S410.
* Jacobs Jr, W. R. (2014). Gene transfer in Mycobacterium tuberculosis: shuttle phasmids to enlightenment. *Microbiology spectrum*, *2*(2).
* Jain, P., Hsu, T., Arai, M., Biermann, K., Thaler, D. S., Nguyen, A., ... & Larsen, M. H. (2014). Specialized transduction designed for precise high-throughput unmarked deletions in Mycobacterium tuberculosis. *MBio*, *5*(3), e01245-14.
* Lin, C. S., Chen, B. Y., Park, T. H., & Lim, H. C. (1998). Characterization of bacteriophage λ Q− mutant for stable and efficient production of recombinant protein in Escherichia coli system. *Biotechnology and bioengineering*, *57*(5), 529-535.
* Marinelli, L. J., Piuri, M., Swigoňová, Z., Balachandran, A., Oldfield, L. M., van Kessel, J. C., & Hatfull, G. F. (2008). BRED: a simple and powerful tool for constructing mutant and recombinant bacteriophage genomes. *PLoS One*, *3*(12), e3957.
* Payne, K., Sun, Q., Sacchettini, J., & Hatfull, G. F. (2009). Mycobacteriophage Lysin B is a novel mycolylarabinogalactan esterase. *Molecular microbiology*, *73*(3), 367-381.
* Tanji, Y., Furukawa, C., Na, S. H., Hijikata, T., Miyanaga, K., & Unno, H. (2004). Escherichia coli detection by GFP-labeled lysozyme-inactivated T4 bacteriophage. *Journal of biotechnology*, *114*(1), 11-20.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_