

A Vector Set For Plasmid-Based RNAi Studies in Mammalian Cells

Version 1.0, May 2004

Iain D.C. Fraser

Alliance for Cell Signaling, Molecular Biology Laboratory
California Institute of Technology, Pasadena, CA

Introduction

RNA interference (RNAi) has recently emerged as a powerful experimental tool in mammalian cell biology. Double-stranded RNA (dsRNA)-induced gene silencing was originally discovered as a mechanism for regulating expression of endogenous genes in *C. elegans* (1), although a similar phenomenon, termed cosuppression, had been demonstrated in plants (2). Studies of the RNAi mechanism demonstrated that the degradation of the target mRNA was mediated by short fragments of dsRNA (3, 4), which were shown to be derived from the longer dsRNA template through the action of the ribonuclease dicer (5). These short interfering RNAs (siRNAs), when exogenously introduced, can enter the RNAi pathway downstream of dicer and efficiently reduce expression of their target gene. The RNAi pathway was not initially considered a useful experimental tool in mammalian cells due to interferon responses to the dsRNA trigger (6). However, Tuschl and colleagues hypothesized that this approach could be applied to mammalian cells if introduction of the shorter siRNAs could avoid activation of the interferon pathway (7). They confirmed their hypothesis in several mammalian cell lines (7), and RNAi has now been shown to be a valuable experimental tool in a variety of biological systems (8).

The direct transfection of chemically synthesized siRNA duplexes into mammalian cells, as originally demonstrated by the Tuschl lab, is currently the most popular approach to RNAi. However, the applicability of this technique is dependent on the transfectability of the model cell system, and since the presence of the siRNA in the cell is transient, longer term experiments are more difficult. This issue has been addressed by the development of approaches that permit the expression of siRNAs from DNA-based plasmid vectors. Several groups have shown that siRNAs can be transcribed as stem-loop hairpin structures under the control of RNA polymerase III (pol III) promoters (9-13). Figure 1 shows a schematic of how such a short hairpin RNA (shRNA) can be transcribed from the promoter for the RNA component of RNase P, H1. Similar expression cassettes have also been developed using the promoter for the small nuclear RNase, U6.

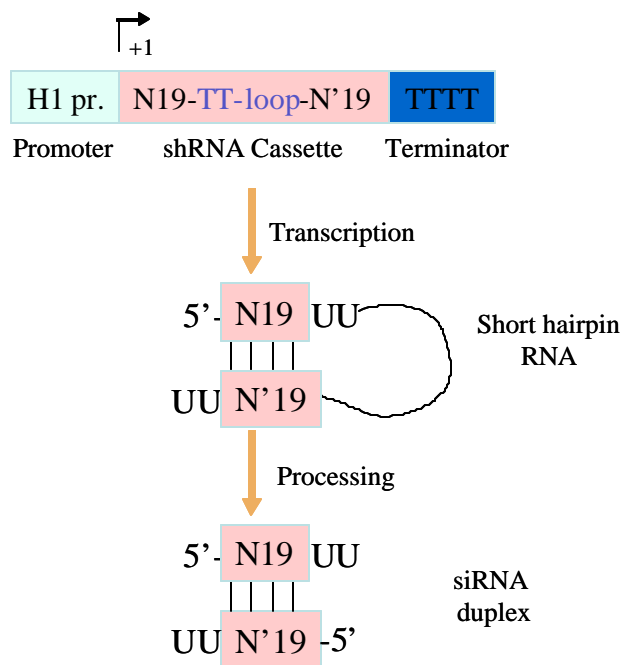


Fig. 1. Expression of an shRNA from a pol III expression cassette. N19 represents a 19 nucleotide sequence homologous to the target gene of interest, and N'19 represents the complementary sequence of the antisense strand. The cassette is designed such that the first base of the N19 sequence is positioned at the transcriptional start site of the H1 promoter. The termination signal for transcription is a sequence of four thymidines.

The pol III promoter-based system described in this document allows for the expression of siRNAs from either mammalian expression vectors or viruses, thus permitting the transduction of difficult-to-transfect cell lines and primary cell cultures. This system also allows for the creation of stable cell lines depleted in a specific target gene.

Based on the architecture of the expression cassette shown in Fig. 1, several groups have developed vector systems that permit the insertion of an shRNA cassette between a promoter and terminator sequence using restriction enzyme sites adjacent to these transcriptional elements (Fig. 2).

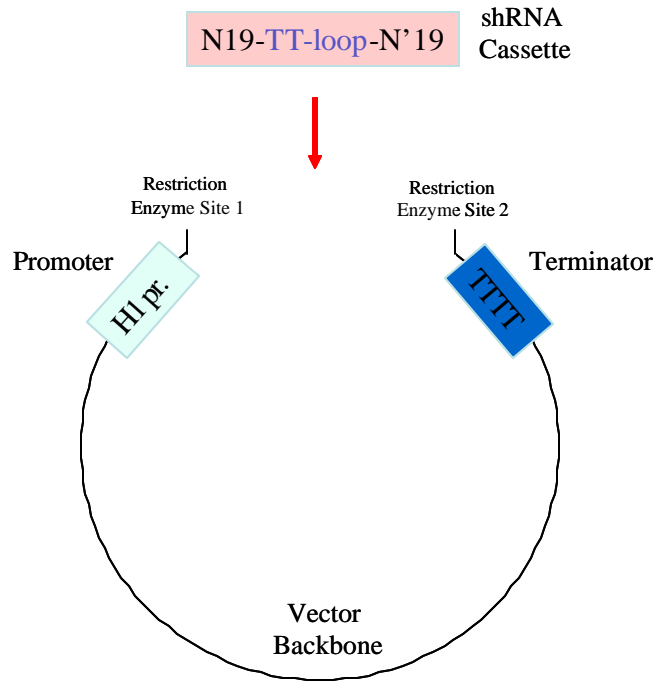


Fig. 2. Creation of a plasmid vector for shRNA expression. The shRNA cassette is created by the annealing of two complementary oligonucleotides containing the desired sequence.

Using this approach, the vector backbone can be varied to permit expression of the shRNA in a variety of contexts. Many companies now sell vectors for shRNA expression based on this principal. Although powerful, this approach requires laborious cloning and sequencing steps for every vector into which a given shRNA is cloned. This document describes a more flexible vector system for shRNA expression.

Gateway-Based Vectors for shRNA Cloning

Background

Using the Gateway cloning technology (Invitrogen), we have developed a more flexible vector system that reduces time spent subcloning shRNAs into multiple expression vectors. Detailed background on Gateway cloning is available at the Invitrogen Web site (<http://www.invitrogen.com>). Briefly, the system uses site-specific recombination instead of traditional restriction enzymes and ligase for subcloning recombinant DNA. Although this system was developed primarily for the subcloning of cDNAs, it can be readily applied to the subcloning of the pol III-shRNA cassettes shown in Fig.1 and Fig. 2.

The first step in the use of the Gateway system is to clone the DNA sequence of interest into a so-called entry vector. The basic property of an entry vector is that it contains the attL1 and attL2 recombination sites at either side of the DNA sequence that will eventually be shuttled into different expression systems. We have created entry vectors containing either a human or mouse H1 promoter (pEN_hH1c and pEN_mH1c, respectively). Figure 3 shows a schematic detailing the shRNA cloning process and the restriction enzyme sites used for these entry vectors.

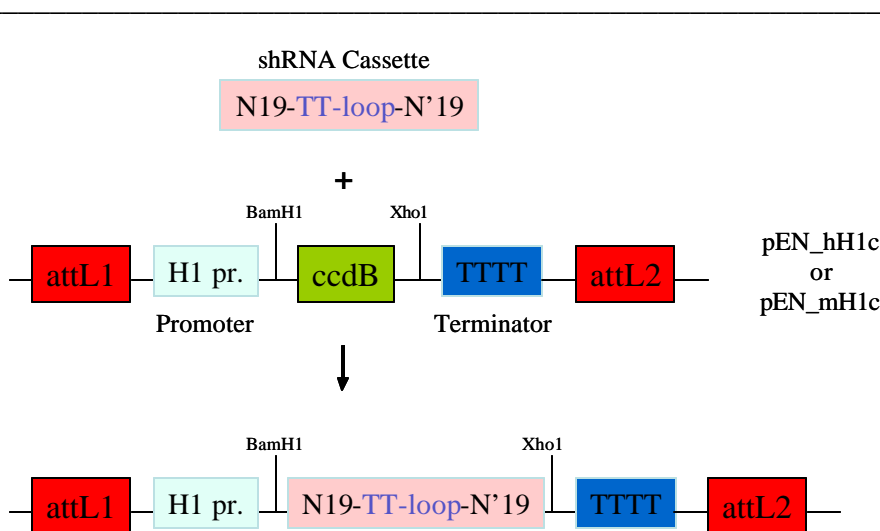


Fig. 3. Ligation of an shRNA cassette into either the pEN_hH1c or pEN_mH1c entry vectors.

Note that full vector maps for the pEN_hH1c and pEN_mH1c plasmids are provided in Appendix 1. Table 1 provides names, ATCC IDs, AfCS bar codes, and relevant vector details for each of these vectors.

Table 1. Details of entry vectors for shRNA cloning.

Vector name	ATCC ID	AfCS bar code	Vector details
pEN_hH1c	10326368	P05EENHH1CXG	shRNA cloning and expression; Human H1 promoter; Entry vector backbone; +ccdB in parent
pEN_mH1c	10326369	P06EENMH1CXG	shRNA cloning and expression; Mouse H1 promoter; Entry vector backbone; +ccdB in parent

The schematic in Fig. 3 shows an element located between the BamH1 and Xho1 enzymes sites in the pEN_H1 vectors labeled ccdB. This element is excised during the cloning of the shRNA cassette. The ccdB gene product is toxic to commonly used strains of *E. coli*, so the presence of this gene in the pEN_H1 parent vectors means that when the products of the shRNA ligation reaction are used to transform a common *E. coli* strain, there can be no background colonies from uncut or single-cut parent. We have found that this approach significantly increases the percentage of transformants containing the desired shRNA sequence.

Design and Synthesis of an shRNA Sequence against a Target Gene

Selection Criteria for Gene-Specific shRNA Sequences

Until recently, the dogma in the RNAi field was that there was no reliable method of rationally designing effective siRNA target sequences for any given mRNA. Our general approach in designing shRNAs for expression in the pEN_H1 vectors has been to choose four sequences, spaced throughout the mRNA of each gene we want to target, and then to test each shRNA for efficacy. Some general rules for sequence selection are as follows:

1. Choose a gene-specific sequence of 19 to 21 nucleotides.
2. The GC content of the sequence should be between 45% and 55%.
3. The T_m of the sequence should be between 45 °C and 65 °C.
4. Avoid runs of two or more As at beginning of sequence (causes premature termination of transcription; see Fig.4).
5. Avoid runs of two or more Ts at end of sequence (causes premature termination of transcription; see Fig.4).
6. Run the sequence through a BLAST search to ensure specificity for the target gene.

More recently, several groups have developed algorithms that permit more rational siRNA design. Dharmacon (<http://www.dharmacon.com>) has developed a process termed SMARTpool, which uses over 30 criteria that improve the likelihood that a selected sequence will represent an effective siRNA. Dharmacon also offers a free basic version of their algorithm (using a handful of their criteria) on their Web site in their siDESIGN center. This does not provide the high success rate of their custom SMARTpool sequence

Cloning and Testing shRNAs

Cloning of shRNA Linkers into the pEN_hH1c or pEN_mH1c Vector

Upon synthesis of sense and antisense oligos for an shRNA cassette, the oligos must be annealed to form a linker for ligation into one of the pEN_H1 entry vectors shown in Table 1. This procedure is detailed in the AfCS protocol PP00000230. Each candidate shRNA clone must be fully sequence verified to ensure that it retains 100% homology to the target gene.

Testing of shRNA Linkers in the pEN_H1 Vectors

The pEN_H1 vectors contain all the elements necessary for shRNA expression in mammalian cells. Once an shRNA has been sequence validated, the resulting construct can be transfected into a mammalian cell by established methodology. The effect of the shRNA on target gene expression can then be assessed by either RT-PCR or Western blot.

Note that in cell lines that are difficult to transfect, it may be difficult to assess shRNA efficacy due to low transfection efficiency of the plasmid. We find that for shRNA testing, it is useful if a full-length cDNA construct is available for the target gene. In these circumstances, we generally cotransfect HEK293 cells with the shRNA-expressing plasmid and another plasmid expressing a GFP-tagged fusion of the target gene. shRNA efficacy is then assessed by Western blot using an antibody against GFP.

In cell lines that transfect to a high efficiency, validated shRNAs transiently expressed from the pEN_H1 vectors may be sufficient for assessing the consequences of target gene knockdown in many cell-based assays. However, in more difficult-to-transfect cell lines and primary cells, an alternative transduction approach is necessary to ensure that an shRNA is expressed throughout a cell population. One such approach would be to express an shRNA from a mammalian expression vector containing a drug resistance gene that would permit selection of a stable cell line after transfection. Another possibility would be to subclone the shRNA cassette into a viral vector, allowing more efficient infection of cells with virus, followed by the subsequent selection of a stable line in cases where the viral vector carries a drug resistance gene.

Use of a Gateway entry vector for shRNA cloning opens up the possibility of simple subcloning of H1 promoter-shRNA cassettes into multiple expression systems.

Subcloning of Pol III-shRNA Cassettes via the LR Recombination Reaction

As mentioned above, the Gateway cloning system can reduce time spent subcloning shRNAs into different expression systems. We have created a wide range of gateway-compatible expression vectors, allowing users to choose the shRNA expression method best suited to their needs.

Figure 3 showed that after an shRNA cassette is ligated into one of the pEN_H1 vectors, the H1 promoter-shRNA-transcription terminator elements are flanked by attL recombination sites. Consequently, if such an attL-containing vector is combined with an attR-containing vector in an LR recombination reaction, the H1 promoter-shRNA-transcription terminator elements will be moved to the attR-containing vector.

Figure 5 shows a schematic for this LR recombination reaction. The reagents and experimental protocol for an LR reaction are available from Invitrogen.

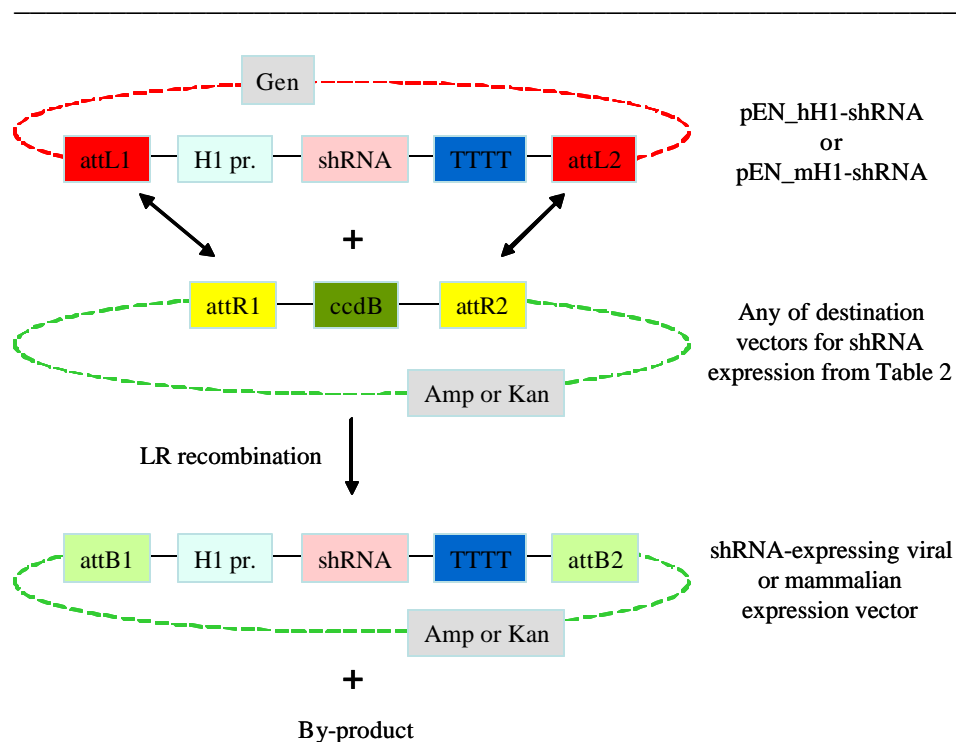


Fig. 5. LR recombination reaction resulting in the movement of an shRNA expression cassette from a pEN_H1 vector to one of the destination vectors developed for shRNA expression.

It should be noted from Fig. 5 that for an entry and destination vector to be compatible in an LR reaction, they must have different bacterial selection markers to allow for appropriate selection of the resulting expression clones. For this reason, we

developed the pEN_hH1c and pEN_mH1c entry vectors as gentamicin-resistant plasmids (see maps in Appendix 1). This allows the shRNA expression cassettes to be moved to expression vectors that are either ampicillin- or kanamycin-resistant.

Any expression vector containing the attR1-ccdB-attR2 elements at an appropriate location for an shRNA cassette is a compatible destination vector for shRNA expression. We have created a range of mammalian expression and viral expression vectors for this purpose.

Expression Systems for shRNAs

Maps of the available destination vectors we have constructed for shRNA expression are provided in Appendix 2. Table 2 provides names, ATCC, AfCS bar codes, vector types, and relevant details for each of these vectors.

Note that the presence of the attR1-ccdB-attR2 elements (which total almost 2kb) in each of these destination vectors makes them inappropriate as empty vector controls for experimental work. Therefore, in Table 2, we also provide details of control vectors lacking the attR1-ccdB-attR2 elements present in the corresponding destination vector. Maps for these control vectors are provided in Appendix 3. For the three mammalian expression vectors at the bottom of Table 2, appropriate empty vector controls are available from commercial sources.

Table 2. Details of vector systems for shRNA expression.

Name	ATCC ID	AfCS bar code	Vector type	Features coexpressed with shRNA
pDSL_hpUGIP	10326373	L11DDLUGIPXA	Lentiviral	Ubi-c promoter-driven GFP-IRES-Puromycin
pL_UGIP	10326372	L10GLUGIP1XA	Lentiviral	Ubi-c promoter-driven GFP-IRES-Puromycin (control for pDSL_hpUGIP)
pDSL_hpUGIH	10326379	L24DDLUGIHXA	Lentiviral	Ubi-c promoter-driven GFP-IRES-Hygromycin
pL_UGIH	10326378	L23GLUGIH1XA	Lentiviral	Ubi-c promoter-driven GFP-IRES-Hygromycin (control for pDSL_hpUGIH)
pDSL_hpUC	10326377	L20DDLHPUCXA	Lentiviral	Ubi-c promoter-driven delta CD4
pL_UC	10326376	L19GLUC001XA	Lentiviral	Ubi-c promoter-driven delta CD4 (control for pDSL_hpUC)
pDSL_hpUP	10326375	L12DDLHPUPXA	Lentiviral	Ubi-c promoter-driven Puromycin
pL_UP	10326374	L17GLUP001XA	Lentiviral	Ubi-c promoter-driven Puromycin (control for pDSL_hpUP)
pDSL_hpUG	10326371	L06DDLHPUGXA	Lentiviral	Ubi-c promoter-driven GFP
pL_UG	10326370	L01GLUG001XA	Lentiviral	Ubi-c promoter-driven GFP (control for pDSL_hpUG)
pDS_CXGhp	10326381	B03D0CXGHPXA	Retroviral	CMV promoter-driven GFP and 5'LTR-driven Puromycin
pCXG	10326380	M0001CXG000A	Retroviral	CMV promoter-driven GFP and 5'LTR-driven Puromycin (control for pDS_CXGhp)
pDS_hpSC	10326382	B88DDSHPSCXA	Mammalian	SV40 promoter-driven delta CD4
pDS_hpCG	10326383	B94DDSHPCGXX	Mammalian	CMV promoter-driven GFP and SV40-driven Neomycin
pDS_hpEY	10326384	B95DDSHPEYXK	Mammalian	EF1 promoter-driven YFP and SV40-driven Neomycin

All of the vectors detailed in Tables 1 and 2 are now available from the American Type Culture Collection (<http://www.atcc.org/>). When ordering, specific vectors should be referenced by their ATCC ID.

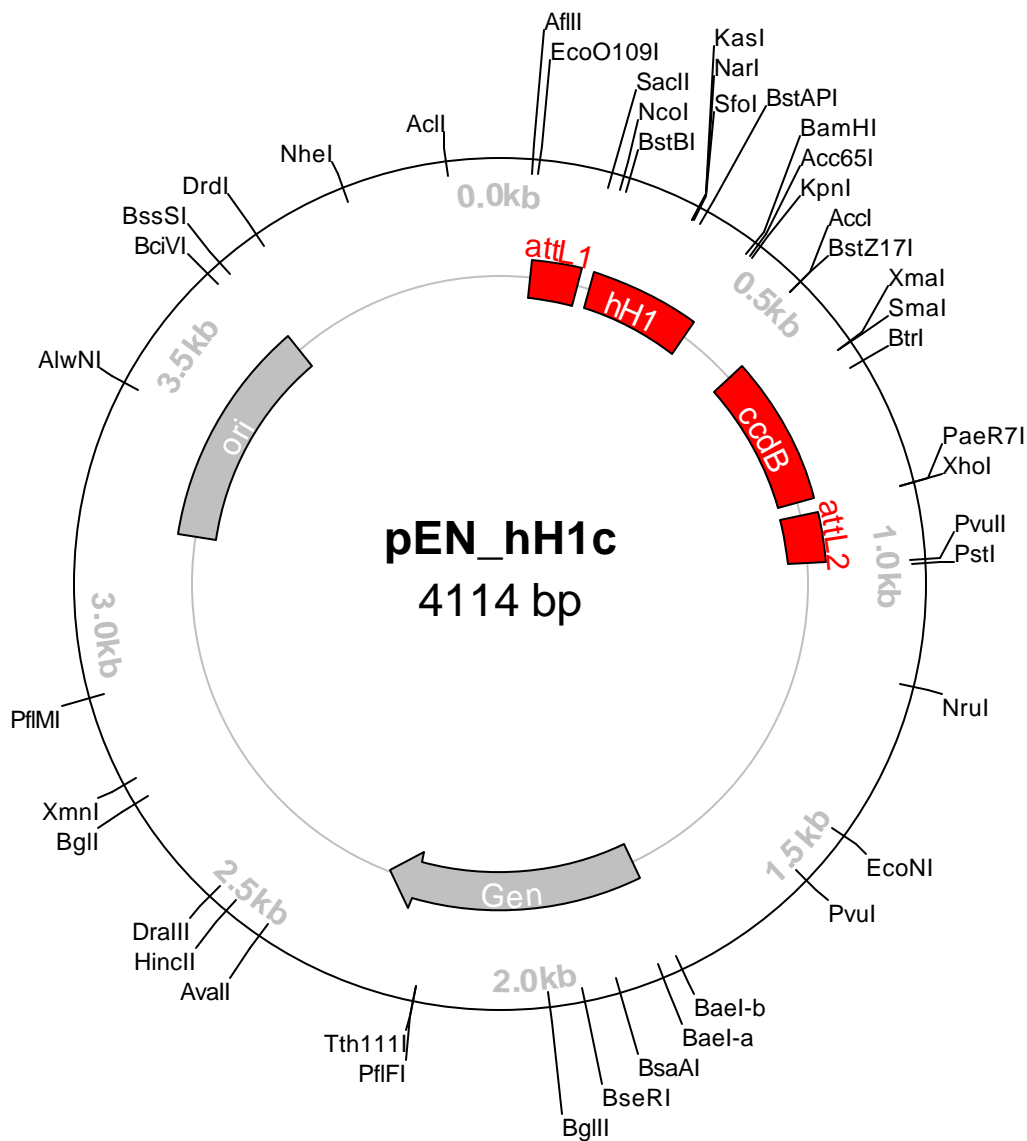
References

1. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, and Mello CC. (1998) *Nature* 391(6669), 806-811. PMID: 9486653.
2. Matzke MA, Aufsatz W, Kanno T, Mette MF, and Matzke AJ. (2002) *Adv. Genet.* 46, 235-275. PMID: 11931226.
3. Hamilton AJ and Baulcombe DC. (1999) *Science* 286(5441), 950-952. PMID: 10542148.
4. Zamore PD, Tuschl T, Sharp PA, and Bartel DP. (2000) *Cell* 101(1), 25-33. PMID: 10778853.
5. Hammond SM, Bernstein E, Beach D, and Hannon GJ. (2000) *Nature* 404(6775), 293-296. PMID: 10749213.
6. Samuel CE. (2001) *Clin. Microbiol. Rev.* 14(4), 778-809. PMID: 11585785.
7. Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, and Tuschl T. (2001) *Nature* 411(6836), 494-498. PMID: 11373684.
8. Hannon GJ. (2002) *Nature* 418(6894), 244-251. PMID: 12110901.
9. Paddison PJ, Caudy AA, Bernstein E, Hannon GJ, and Conklin DS. (2002) *Genes Dev.* 16(8), 948-958. PMID: 11959843.
10. Brummelkamp TR, Bernards R, and Agami R. (2002) *Science* 296(5567), 550-553. PMID: 11910072.
11. Paul CP, Good PD, Winer I, and Engelke DR. (2002) *Nat. Biotechnol.* 20(5), 505-508. PMID: 11981566.
12. Sui G, Soohoo C, Affar el B, et al. (2002) *Proc. Natl. Acad. Sci. U. S. A.* 99(8), 5515-5520. PMID: 11960009.
13. Yu JY, DeRuiter SL, and Turner DL. (2002) *Proc. Natl. Acad. Sci. U. S. A.* 99(9), 6047-6052. PMID: 11972060.

Appendix 1: Vector Maps for pEN_H1 Entry Vectors

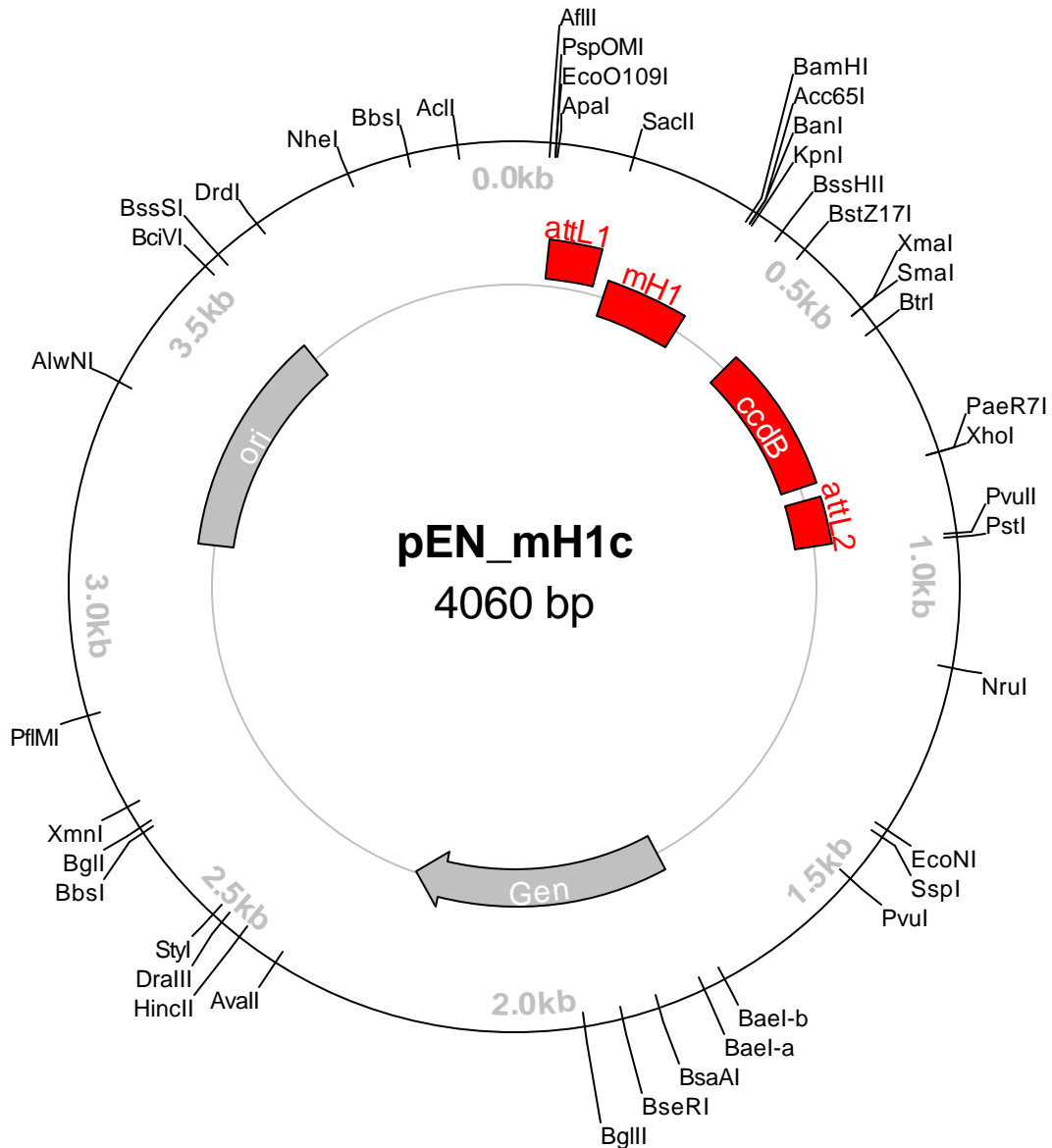
P05EENHH1CXG

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



P06EENMH1CXG

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.

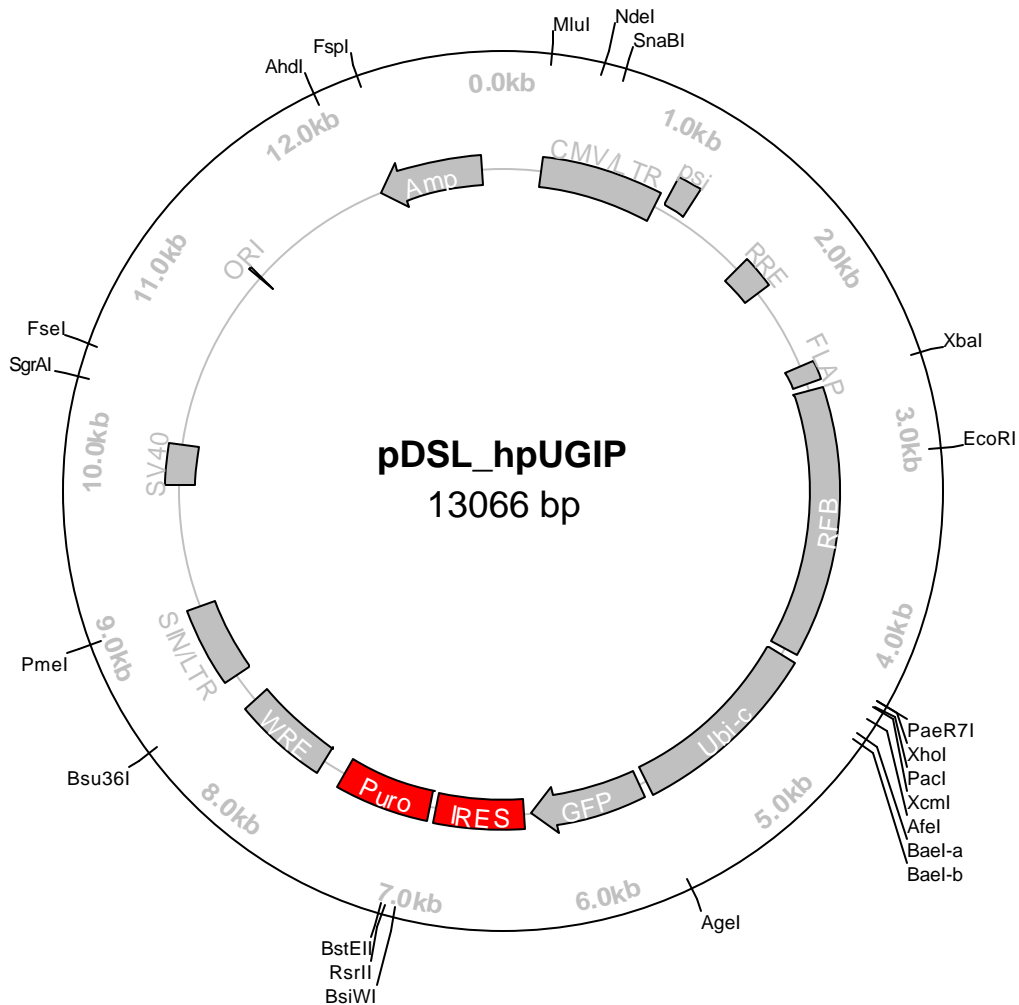


Appendix 2: Vector Maps for Destination Vectors Suitable for shRNA Expression

Note: in all vectors shown in Appendix 2, the RFB feature contains the attR1-ccdB-attR2 elements that define an expression vector as a gateway-compatible destination vector.

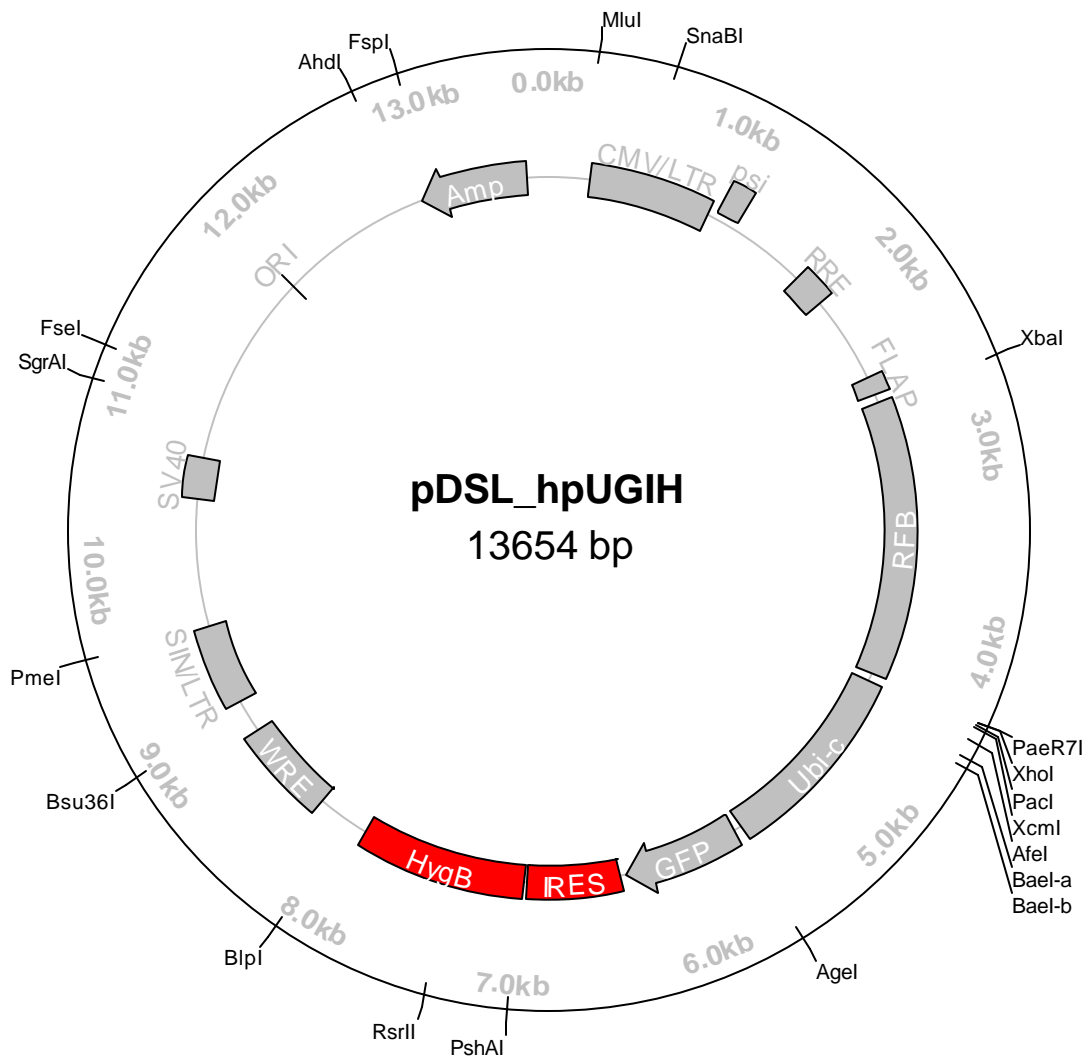
L11DDLUGIPXA

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



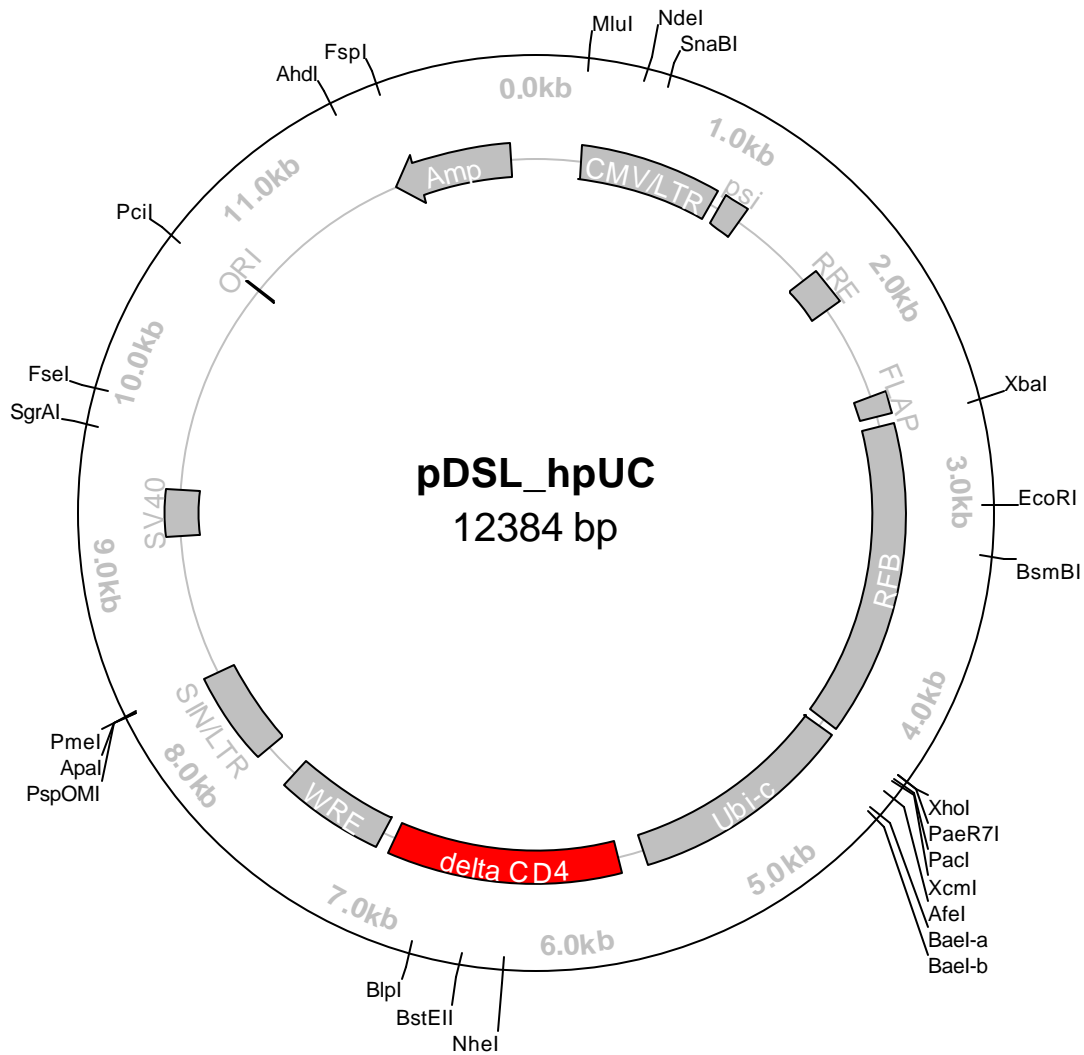
L24DDLUGIHX

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



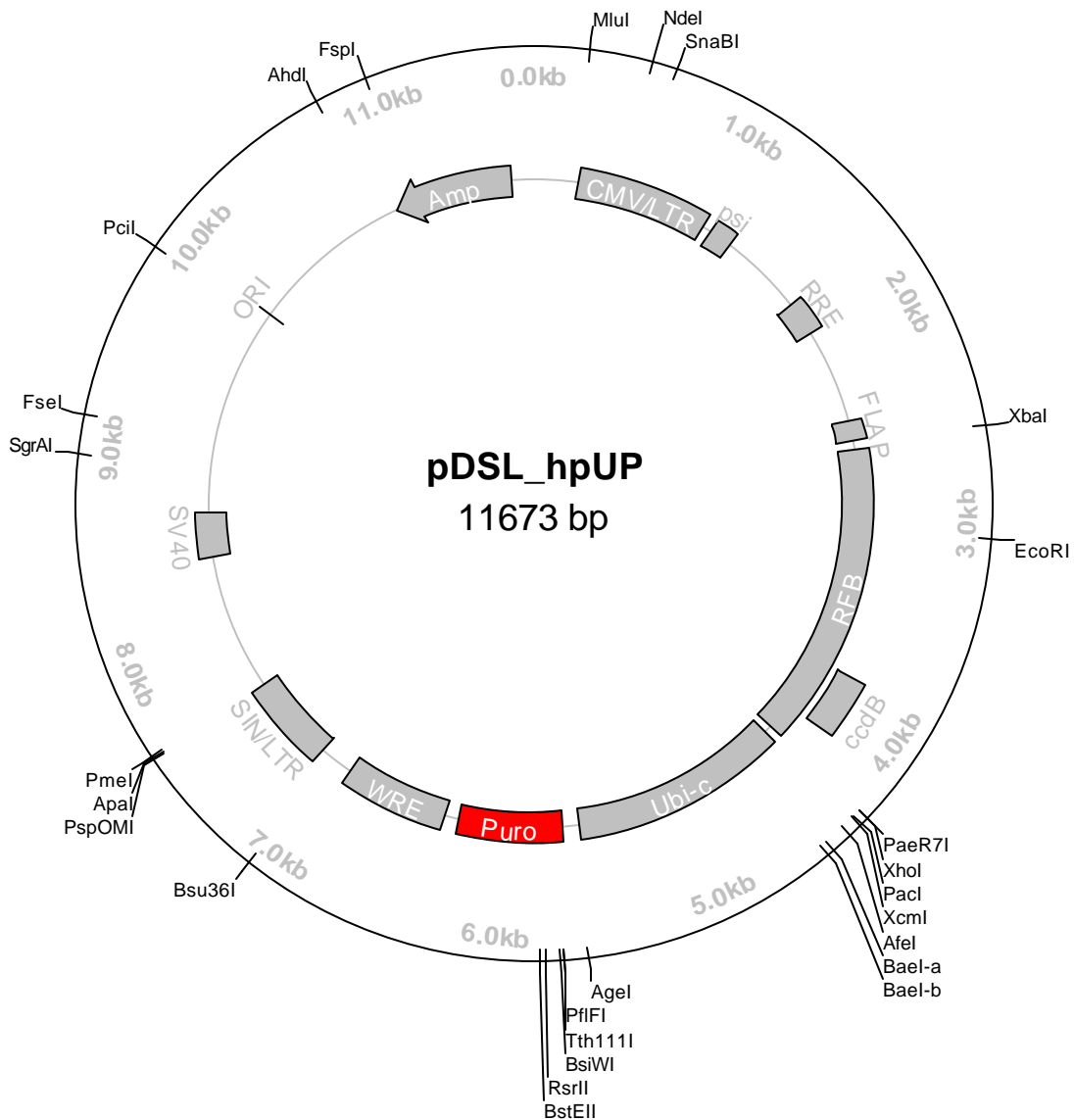
L20DDLHPUCXA

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



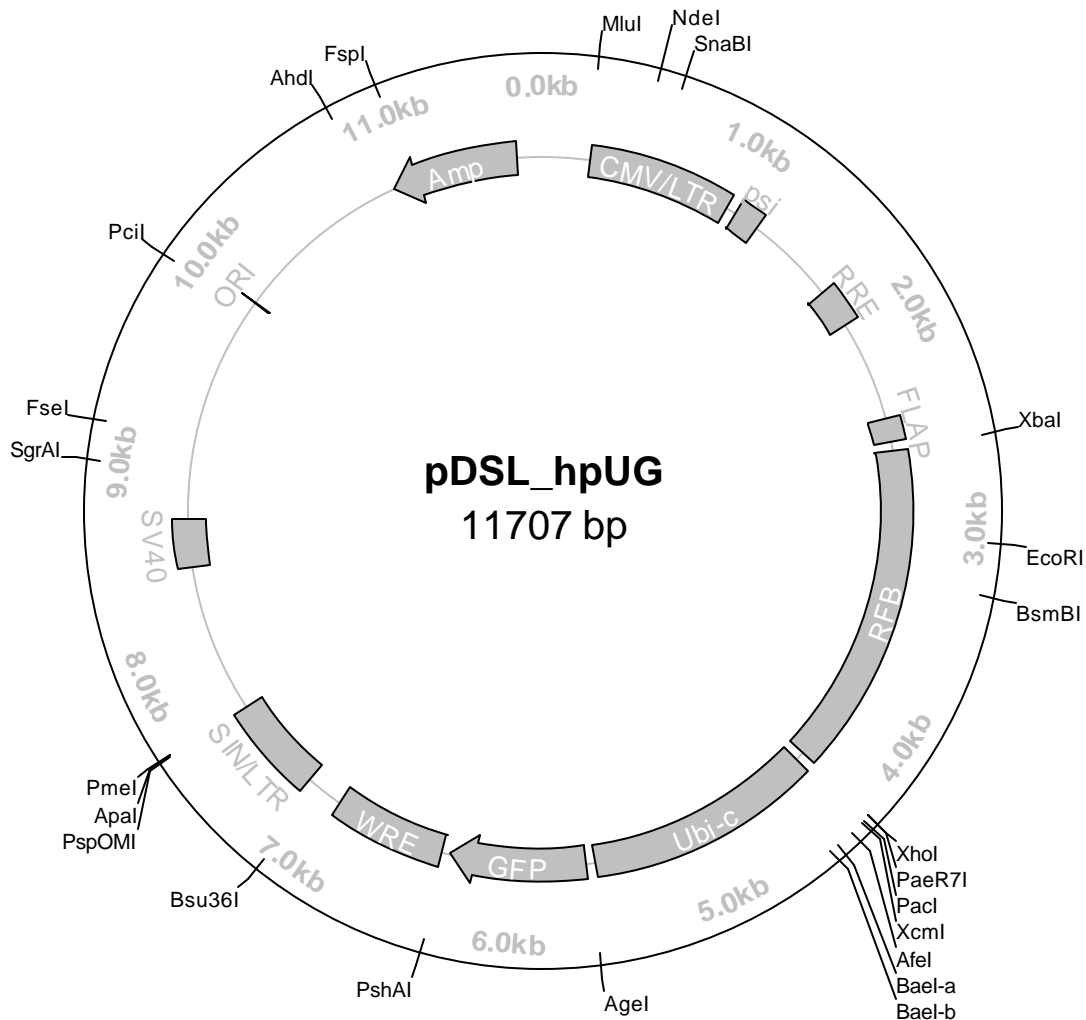
L12DDLHPUPXA

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



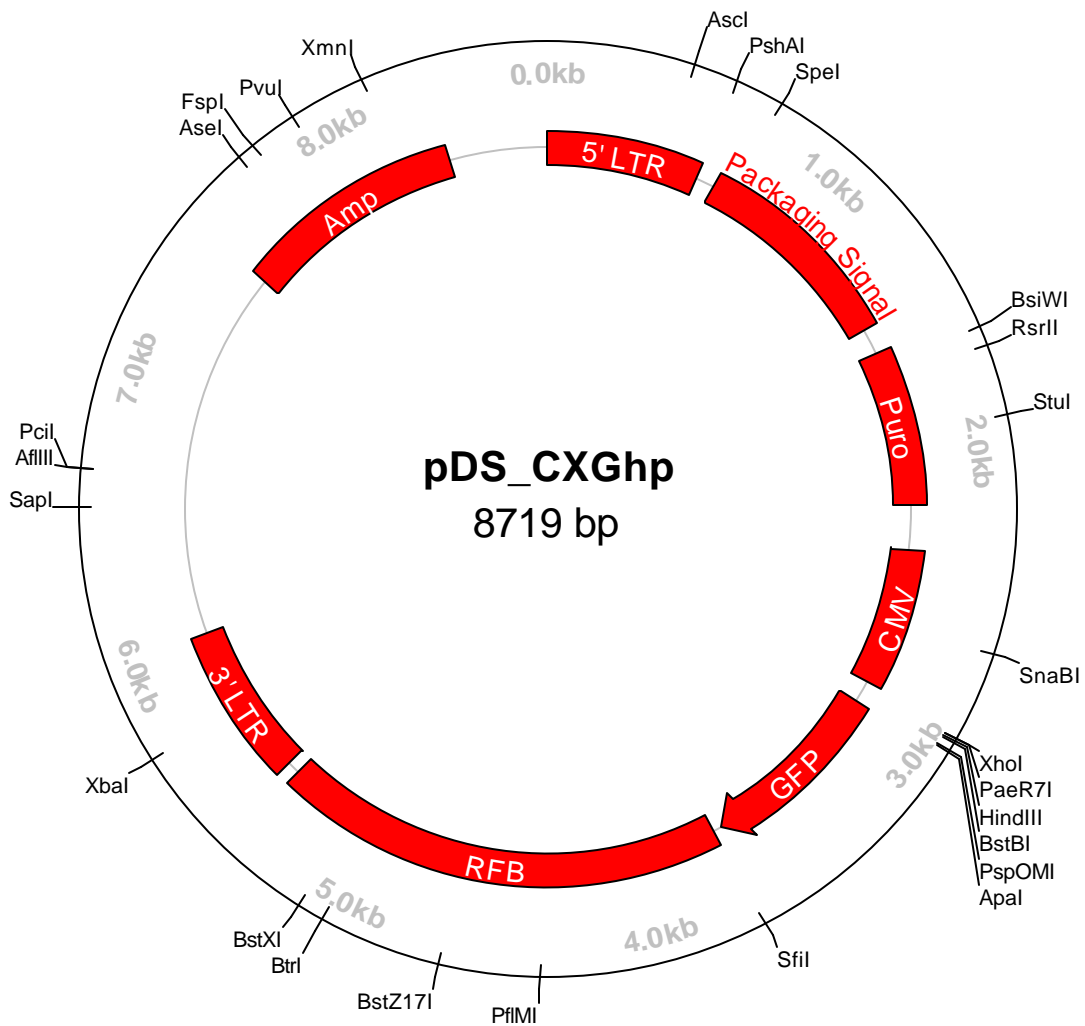
L06DDLHPUGXA

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



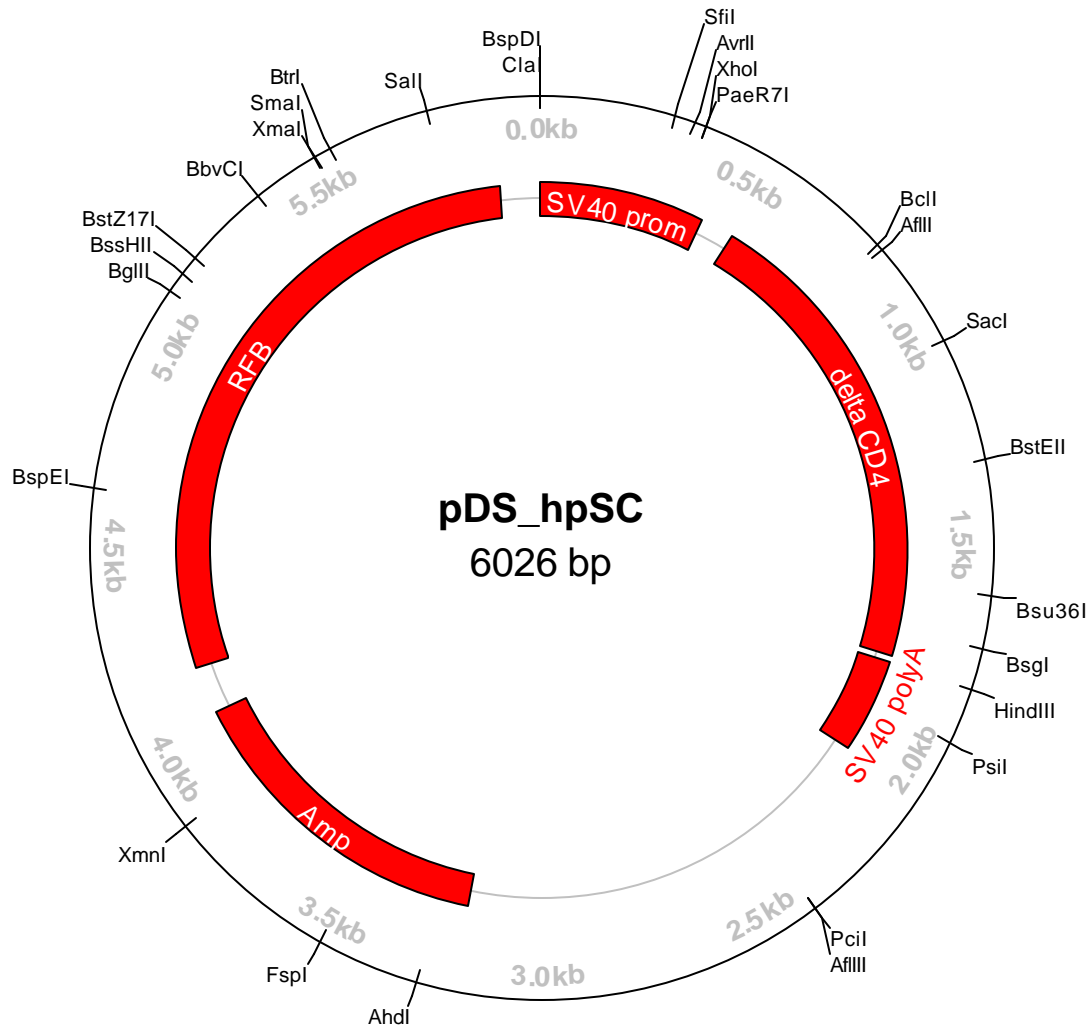
B03D0CXGHPXA

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



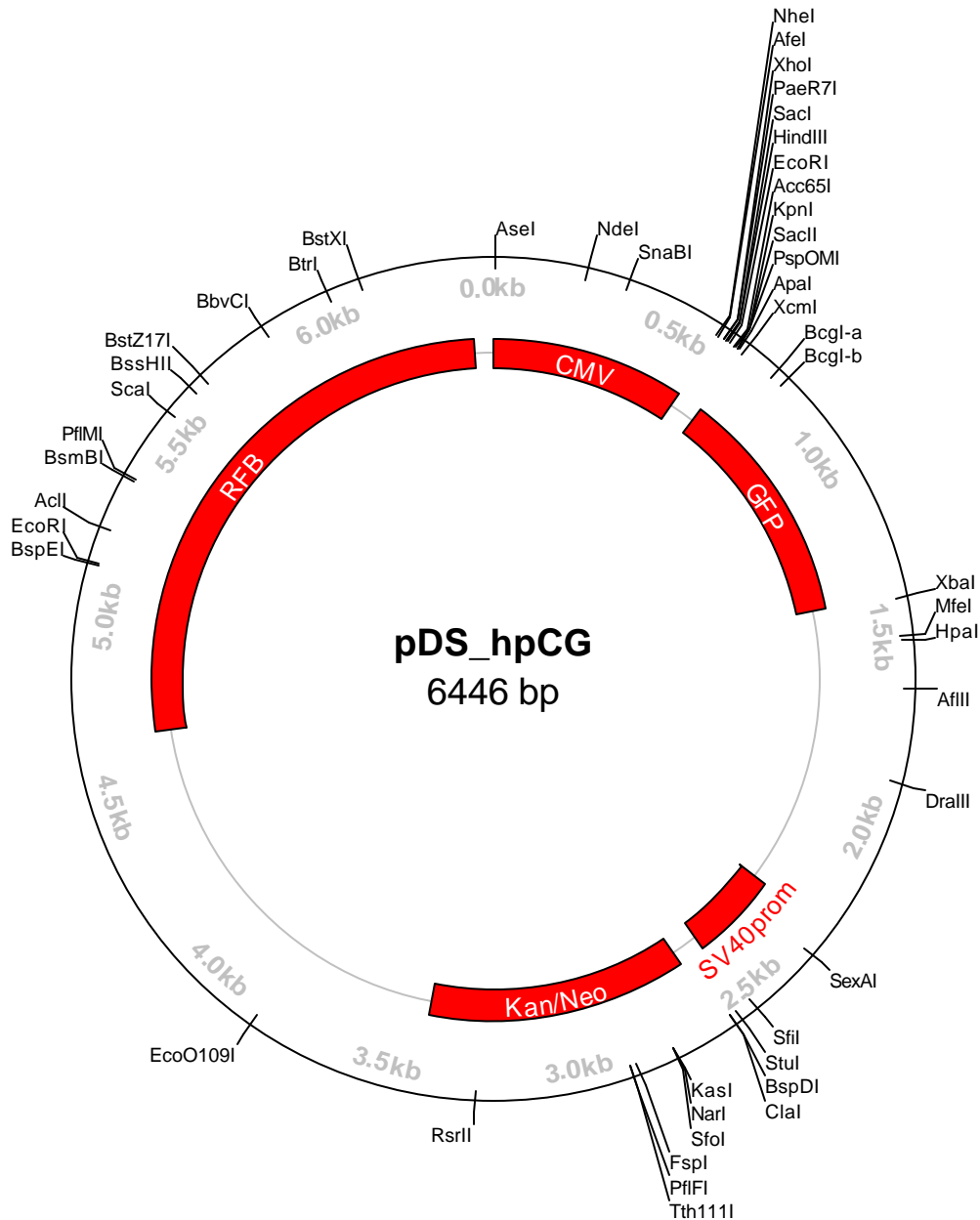
B88DDSHPCXA

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



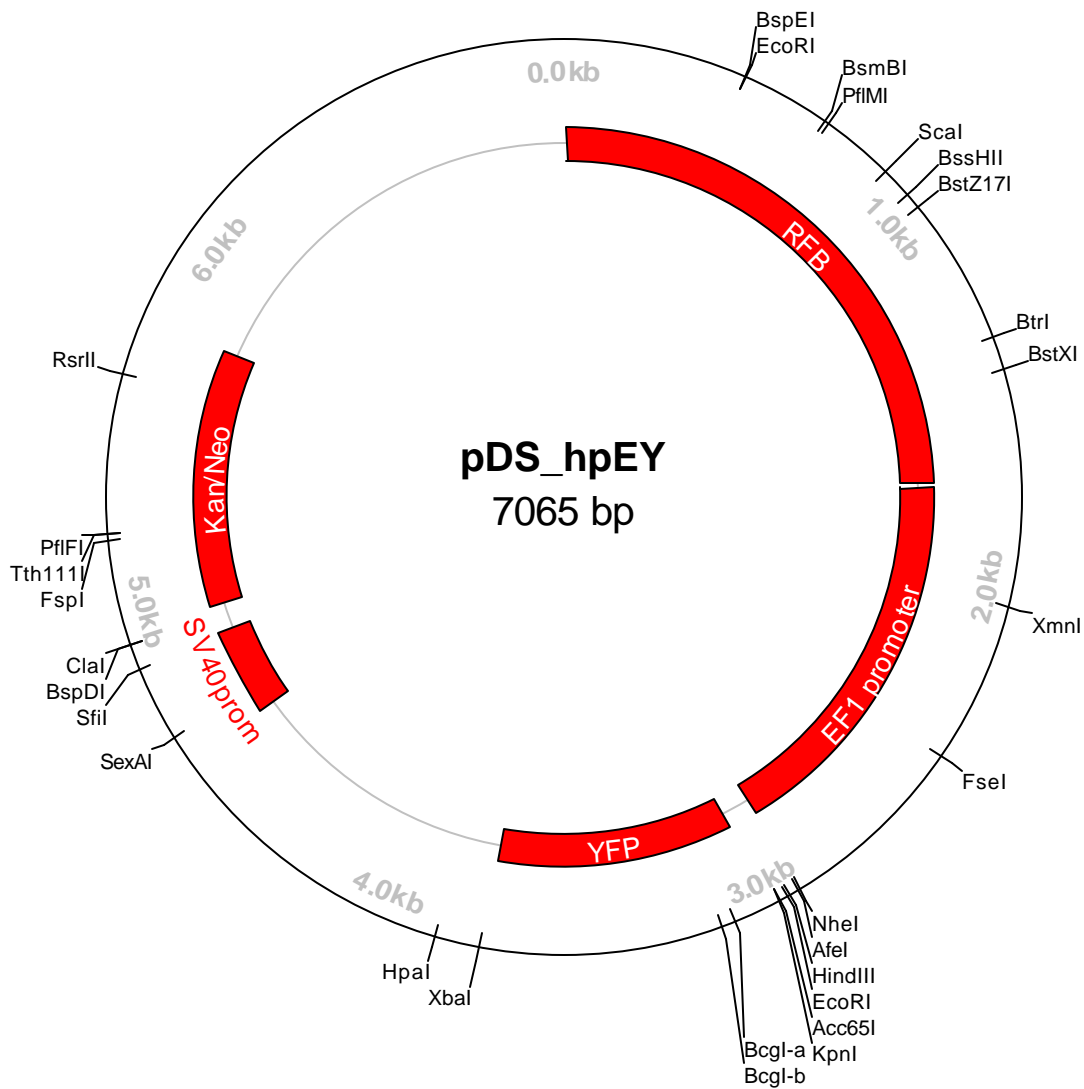
B94DDSHPCGXK

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



B95DDSHPEYXK

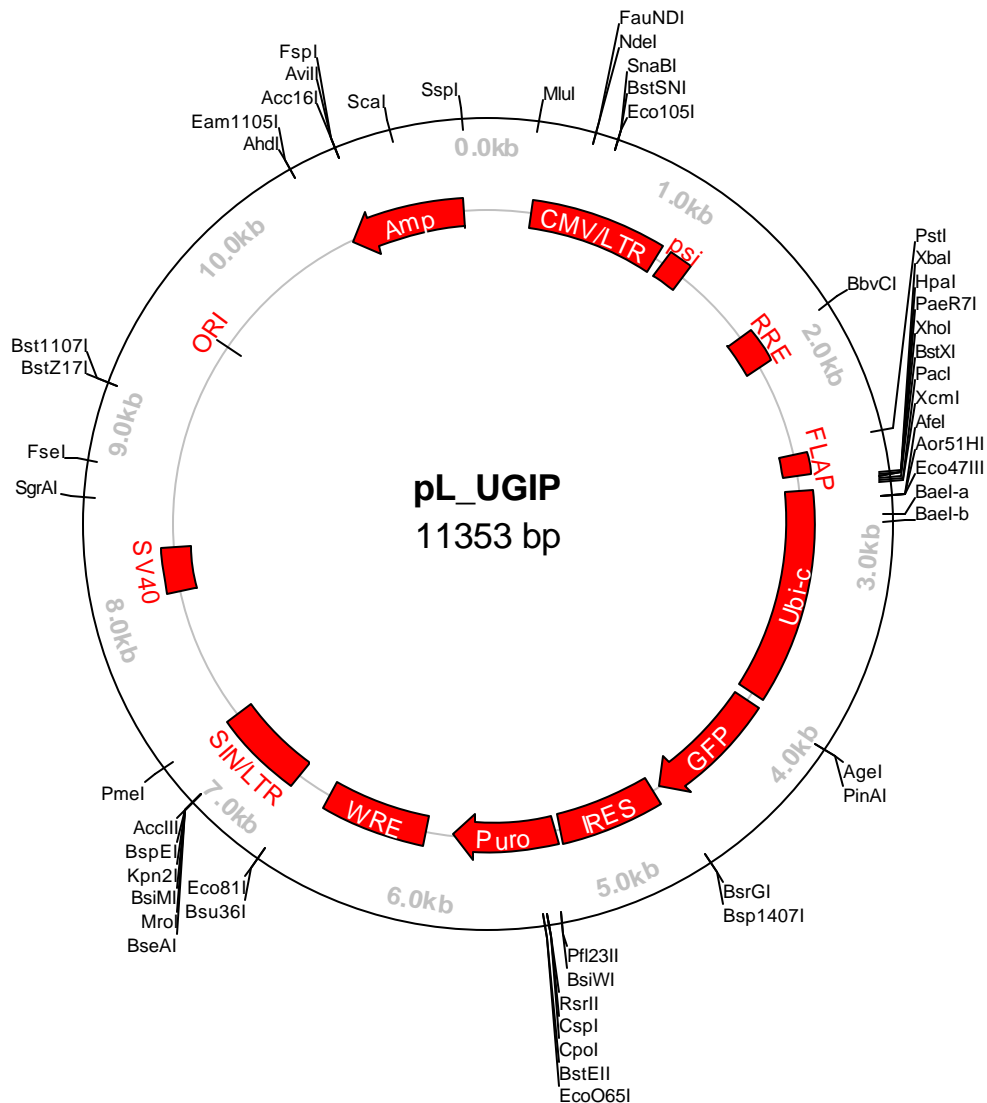
Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



Appendix 3: Maps for Empty Vector Control Plasmids for Use with shRNA-Expressing Destination Vectors

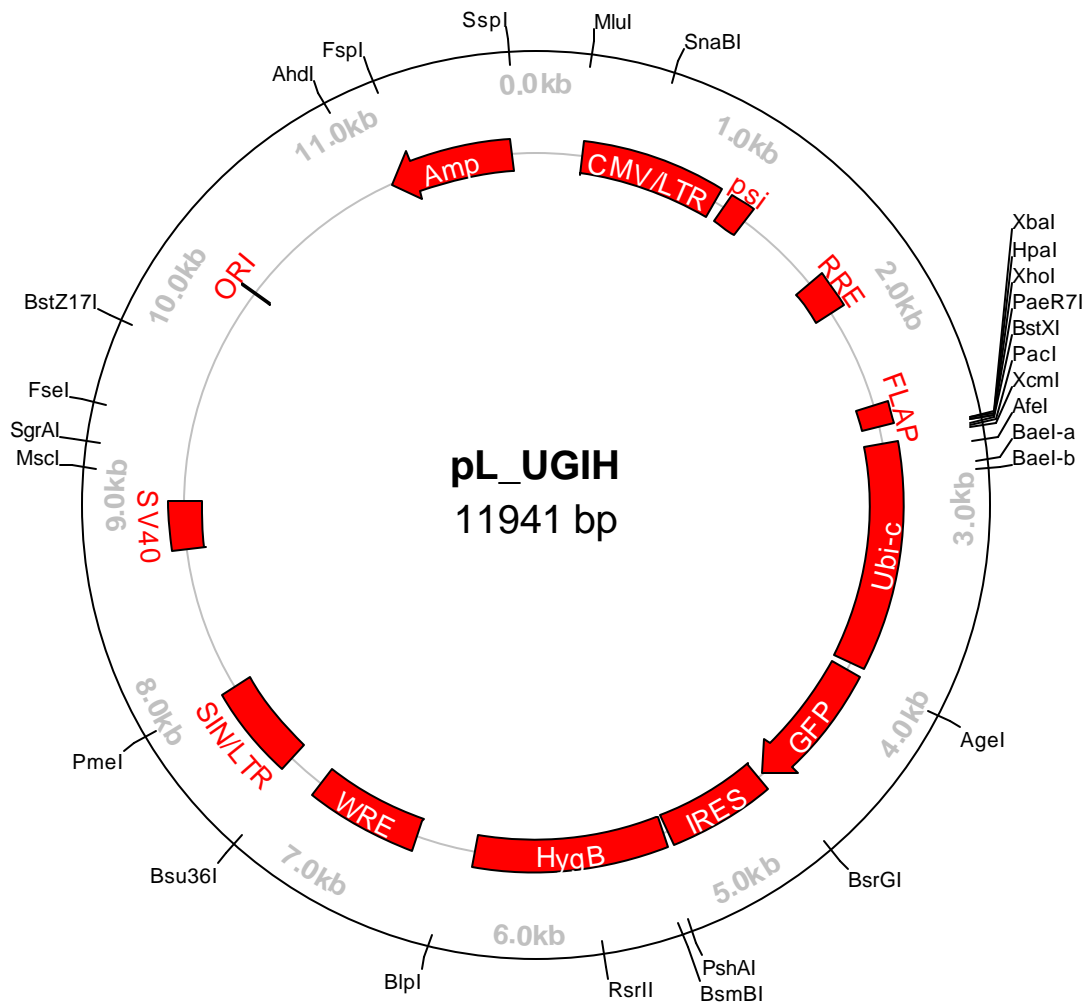
L10GLUGIPIXA

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



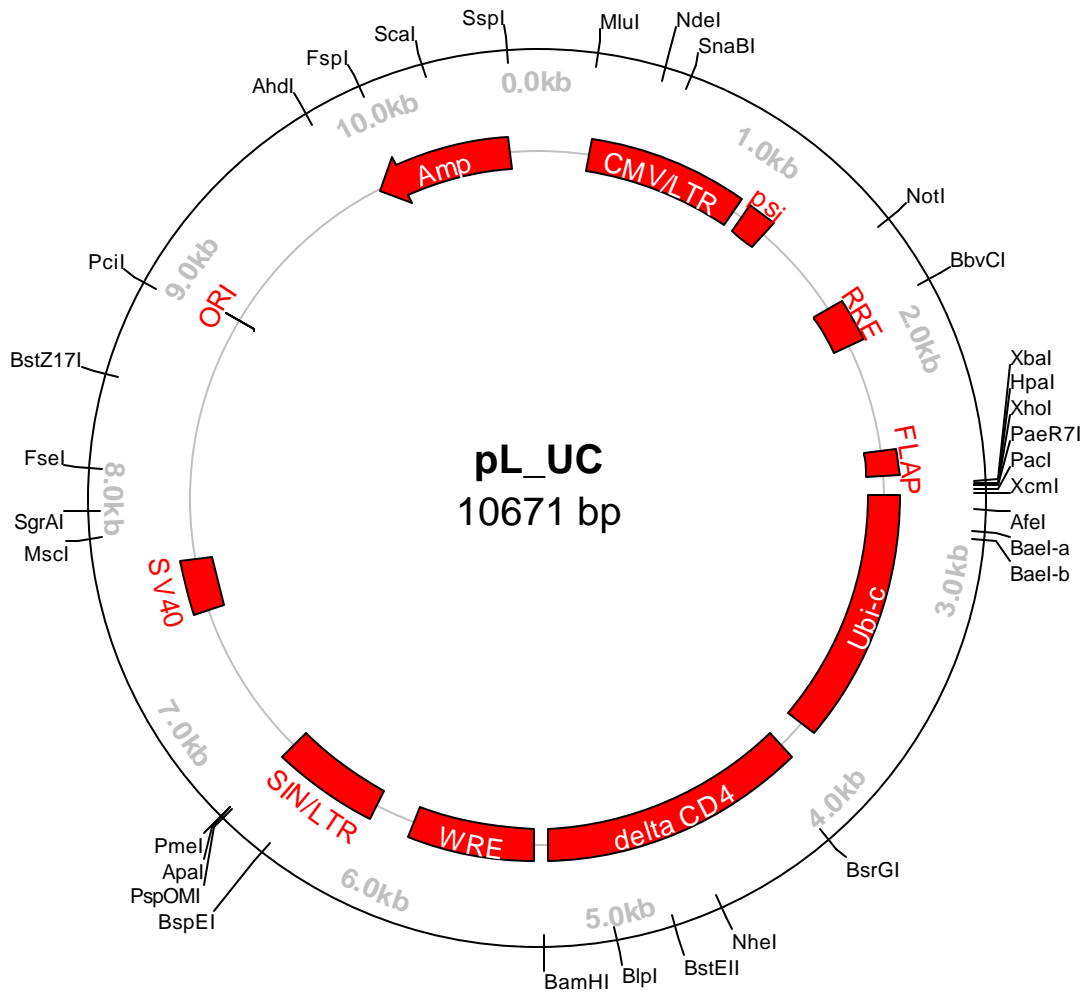
L23GLUGIH1XA

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



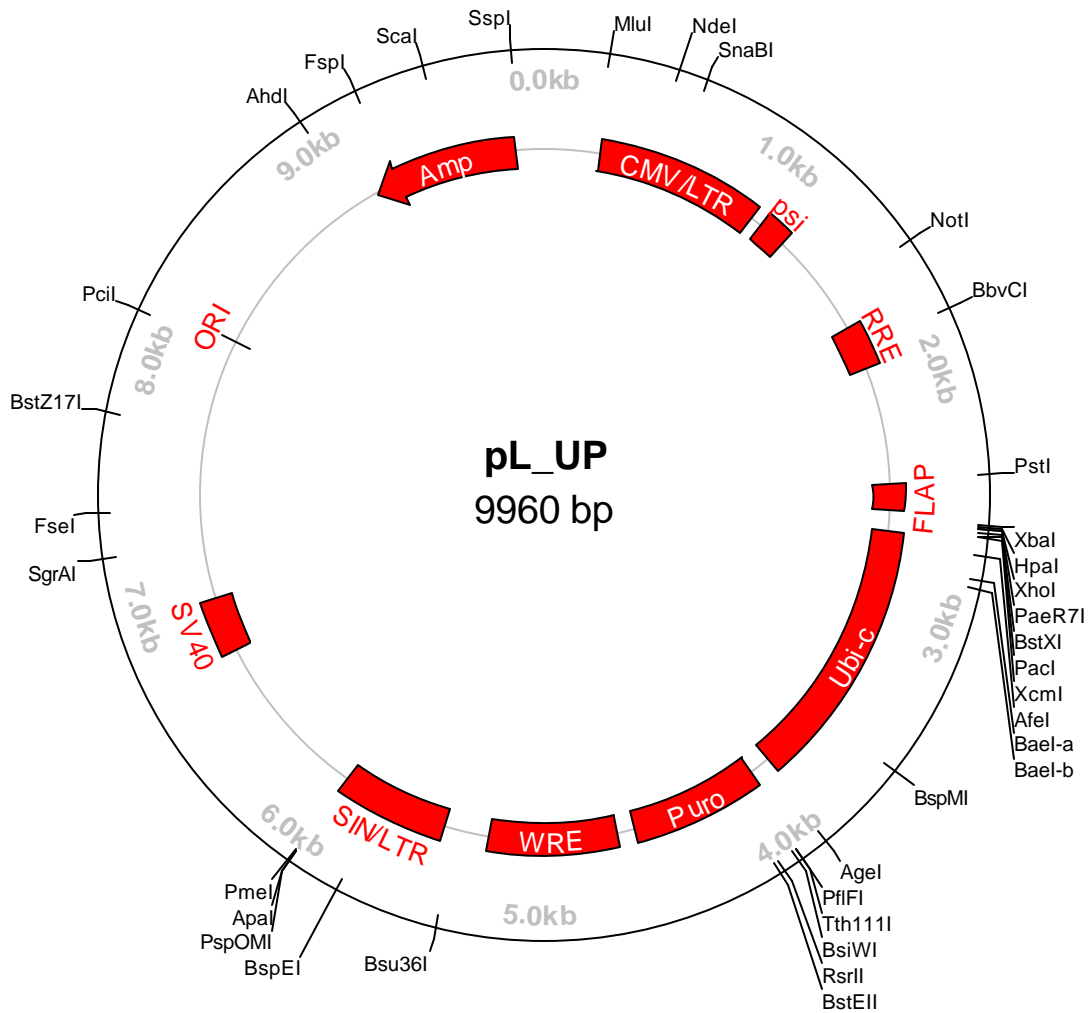
L19GLUC001XA

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



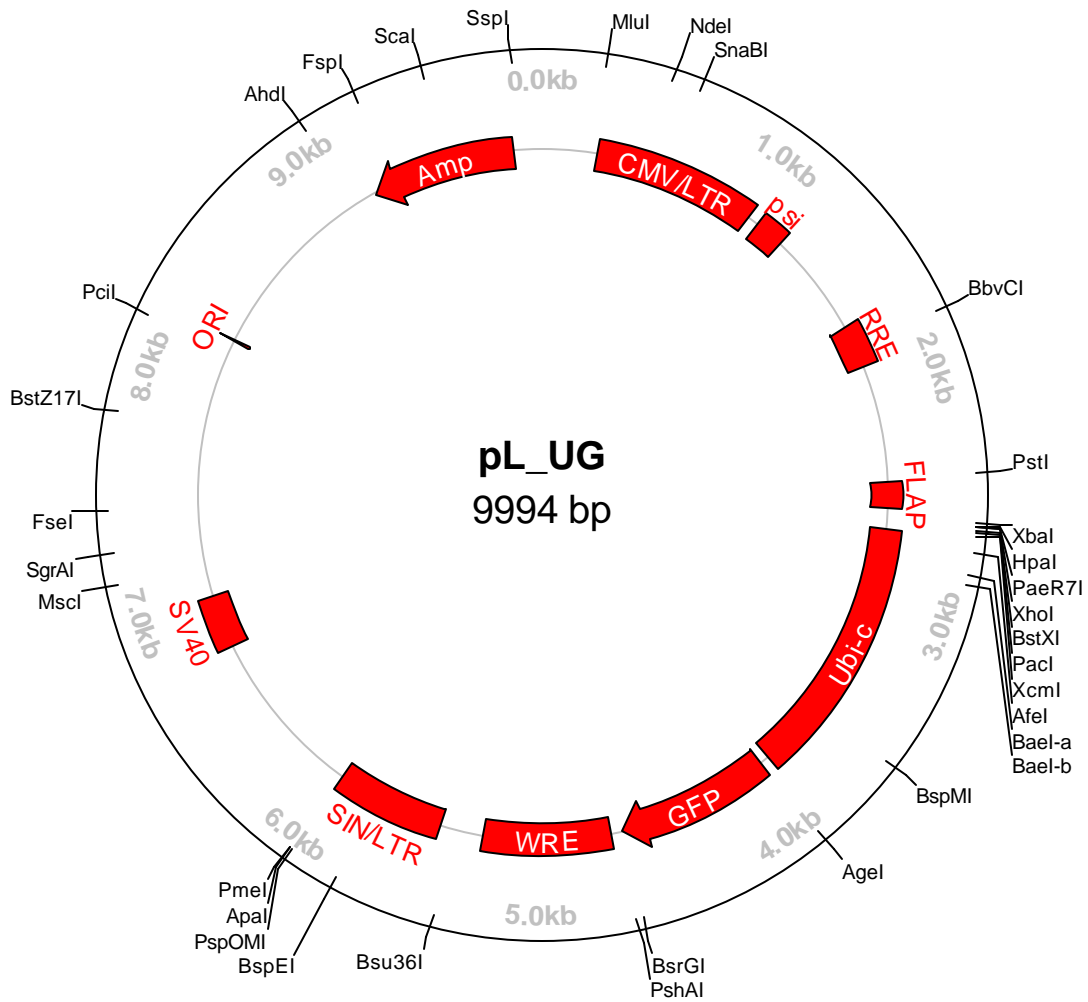
L17GLUP001XA

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



L01GLUG001XA

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.



M0001CXG000A

Full sequence for this plasmid is available at the AfCS data center (<http://www.signaling-Gateway.org/data/Data.html>). Follow the **Plasmid Database: Web interface** link in the **Resources** section and use the above bar code to retrieve plasmid data.

