Mini Review

Beyond pUC: Vectors for Cloning Unstable DNA *

Ronald Godiska, Melodee Patterson, Tom Schoenfeld, and David A. Mead
Lucigen Corporation, Middleton, WI 53562

The foundation of genomic sequence analysis is large-scale cloning and sequencing from shotgun plasmid libraries. Phenomenal advances in the technology for obtaining and assembling sequence information from vast numbers of clones have made this approach practical and efficient. In stark contrast, there has been very little improvement in the vectors used for generating plasmid libraries for shotgun sequencing. Common vectors are typically maintained at high copy number and induce transcription and translation of inserted fragments, causing instability of certain classes of DNA sequences. This “unstable” DNA may result in sequence stacking, clone gaps, or severe difficulties in creating plasmid libraries, especially from DNA with a high percentage (>70%) of adenine and thymine bases (AT-rich DNA).

By far the most commonly used plasmid for constructing shotgun libraries is pUC18 or its closely related derivatives. These vectors have several features widely accepted as advantageous, such as their blue/white screening capacity, large multiple cloning site, and high copy number, as well as the ability to generate RNA transcripts from bacteriophage promoters and single-stranded DNA (ssDNA) from the M13 origin of replication (reviewed in 1). However, very little attention has been devoted to characterizing the potential disadvantages of these traits. For example, specific fragments or even large portions of genomes are recalcitrant to cloning in these vectors, hindering most sequencing efforts. As a result, despite extensive finishing efforts applied over the past years, gaps remain in the genomic sequence of nearly every multicellular organism studied, including those assemblies described as “complete” (2).

Drawbacks of pUC Vectors

Gaps in shotgun libraries and seemingly “unclonable” DNA fragments are quite common. Such DNA is typically characterized by high AT content, strong secondary structure, open reading frames, or *cis*-acting functions (e.g., transcriptional promoters or replication origins). In some cases, most notably AT-rich DNA, the reasons for difficulty in cloning are not well defined. In other instances, one or more features of pUC plasmids have been shown to be incompatible with cloning or stable maintenance of the inserts. Difficulties caused by each of these features are summarized in Table 1.

---

* Excerpted from *DNA Sequencing: Optimizing the Process and Analysis* (2005), J. Kieleczawa (ed.), Jones and Bartlett Publishers, Sudbury MA. www.Jbpub.com. Reprinted with permission.

Table 1. Advantages and disadvantages of pUC-type plasmids

Feature	Advantage	Disadvantage	Targets Selected Against
Blue/white screen or direct selection	Easy screening	Transcription/translation of inserts, false positives, false negatives, or loss of recombinant clones	Toxic ORFs, repeats, AT-rich DNA, short inserts, promoters
High copy number	High plasmid yield	Instability	Large inserts, AT-rich DNA
Ampicillin resistance	Common, inexpensive	Satellite colonies, loss of selection in liquid culture	Slow growers, unstable inserts
M13 origin	Generates ssDNA	Instability when transcribed	Promoters, cDNAs

The central feature of pUC plasmids, which has become ubiquitous among cloning vectors, is the "blue/white" colony screen to detect recombinant plasmids (3). This screen is based upon inactivation of the *lac*Za peptide of beta-galactosidase, which is expressed by the vector. A similar approach is used with direct selection vectors (e.g., pZeRO™; Invitrogen, Carlsbad, CA), except that a gene encoding a lethal product, such as *ccd*B or *sac*B, is used in place of, or in addition to, *lac*Za to select against non-recombinants. Although these screens provide a simple and powerful method to identify most recombinant colonies, they may in fact select against a relatively large number of clones. Both screens induce a high level of transcription and translation of the indicator gene, driven by a promoter in the vector region adjacent to the cloning site. Sequences inserted into the multiple cloning site are also subject to this transcription and translation, which may interfere with cloning of several classes of inserts

Recently, several alternative vectors have been introduced by Lucigen to reduce the cloning bias inherent to pUC plasmids. Dramatic improvements in library construction are possible with these new vectors that reduce the cloning bias. For example, telomeric repeats and other AT-rich fragments from *Pneumocystis carinii* (65% AT) were stable in a low copy, transcription-free vector but were unstable in pUC19. Likewise, the AT-rich genome of *Lactobacillus helveticus* (65% AT) was cloned with 25-fold greater efficiency and significantly less bias using a transcription-free vector. Toxic regions of the mouse hepatitis virus genome were readily cloned and stable in the new type of vector, but they were deleted, rearranged, or slow-growing in conventional vectors. Although successful genomic sequencing with pUC-based vectors has validated their utility and reinforces their continued use, employing alternative vectors may substantially reduce the effort required for genomic sequencing and assembly.

pSMART® Vectors

The pSMART series of vectors (Figure 1, next page) lack several characteristics of pUC18 that have been associated with cloning bias and deletion of inserts. The most notable feature of the pSMART vectors is the absence of transcription into or out of the inserted DNA.

continued, next page

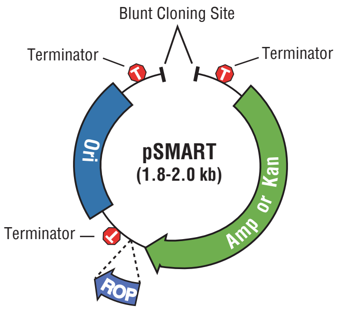


Figure 1. Schematic diagram of pSMART vectors. The pSMART vectors (Lucigen) contain either ampicillin or kanamycin selection. The ROP gene (Repressor of Primer) is present in the low copy versions to decrease copy number. Transcriptional terminators flank the cloning site to prevent insert-driven transcripts from entering the vector. Another terminator follows the drug selection marker to prevent vector-driven transcription of inserts. Vector sizes are 1.8 to 2.0 kb. The pSMART vectors accommodate inserts up to 20 kb.

Vector driven transcription into the insert is eliminated by removal of the *lac*Z promoter and the *lac*Za gene. In addition, a transcriptional terminator downstream of the drug resistance gene prevents read-through of insert sequences from this promoter. Because there is no marker gene to indicate which clones have inserts, the pSMART vectors are supplied by Lucigen pre-cut and dephosphorylated, such that the background of self-ligation is less than 0.1% in most applications. With this low level of background, screening for recombinants is typically not required. The low level of background also facilitates cloning nanogram amounts of sample or cloning samples that produce few colonies for other reasons. Variants of pSMART employ either kanamycin or ampicillin selection, and they contain high or low copy origins of replication.

Extremely efficient cloning in the absence of a visual screen has been verified by sequence analysis of numerous libraries constructed with pSMART. For example, sequencing 32,000 randomly picked clones from a library of 2.5 to 4.5 kb inserts revealed only 40 empty vector clones, or 99.9% recombinants (P. Wilson, Australian Genome Research Facility; personal communication). This level of efficiency is much higher than that obtained with plasmids employing the blue/white screen or *ccd*B selection, which may yield 5% to 10% empty vector clones (i.e., false positives) after screening or selection for recombinants (B. Chiapelli, Washington University Genome Sequencing Center; personal communication).

As expected, the lack of vector-driven expression of insert DNA alleviates the selection against cloning toxic coding sequences, strong promoters, and other sequences that are unstable in pUC-based vectors. A 350-bp fragment encoding a prokaryotic RNase was recovered in either orientation in pSMART, whereas it could be obtained only in the "reverse" orientation in pUC19 (Table 2). Transcription initiated by promoters within the insert is blocked by terminators on either side of the cloning site in pSMART (Figure 1), thereby insulating essential portions of the vector (i.e., the replication origin and drug resistance gene) from transcriptional interference. The efficacy of this arrangement was demonstrated by cloning a 450-bp fragment containing the strong lambda P_R promoter into pSMART and pUC19. Clones containing this promoter were isolated readily in pSMART, whereas they were rare and nearly obscured by empty vector clones when ligated into a dephosphorylated preparation of pUC19 (Table 2).

Table 2. Cloning efficiency of a toxic gene or strong promoter in pSMART or pUC19.

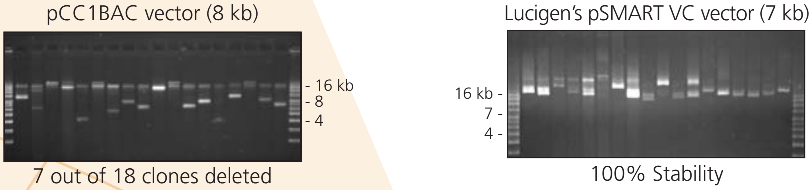
Vector	**Target: RNase Gene** Total cfu (% Forward)	**Target: Lambda P_R** Total cfu (% Intact)
pSMART-HC	45,000 (38%)	8,500 (75%)
pSMART-LC	20,000 (33%)	3,500 (75%)
pUC19	50,000 (0%)	Blue 100,000 (n.d.) White 100 (25%)

Fragments containing a prokaryotic RNase gene or the lambda P_R promoter were amplified with Vent® DNA polymerase (New England Biolabs) and treated with T4 polynucleotide kinase (New England Biolabs), according to the manufacturer's recommendations. Fifty nanograms of each fragment was ligated to blunt, dephosphorylated vector preparations and transformed into electrocompetent *E. coli* cells, using standard molecular biology techniques. Values represent total number of colonies per ligation. PCR was used to determine insert size and percentage of clones in the each orientation.

Several fragments encoding "toxic" regions of the mouse hepatitis genome also were readily cloned and highly stable in pSMART. These same fragments were very susceptible to rearrangement in standard vectors (4). The lack of transcription is likely responsible for producing more representative cDNA libraries as well (R. Drinkwater, Xenome Corp.; personal communication). Cloning into pSMART vectors is expected to increase the stability of trinucleotide repeats, based on their stability in a transcription-free derivative of pUC (5). Libraries from AT-rich genomic DNA tend to be much larger and more representative when constructed in pSMART rather than pUC19. For example, a cosmid containing genomic DNA from *Pneumocystis carinii* (60%–65% AT) was sheared, end-repaired, and size-selected to 2 to 4 kb. Aliquots of this material were cloned into pSMART HCKan or pUC19. No deleted inserts or empty vectors were detected among 151 pSMART clones analyzed. In contrast, over 25% of the clones in the pUC19 library suffered obvious deletions, resulting in plasmids smaller than the parental pUC19 (Figure 2).

pSMART® VC Vectors

The pSMART VC vector (Lucigen) appears to be the most stable vector for cloning extremely recalcitrant DNA fragments, especially those that are over 10 kb in length. This vector incorporates the transcription-free aspects of pSMART into a minimized derivative of pBELOBAC (Figure 3). The pSMART VC vector is single copy for increased cloning stability, but it contains the inducible OriV origin (6) to increase the copy number to 20–100 per cell for DNA isolation. This vector facilitated construction of a stable genomic library of 10–20 kb inserts from *Tetrahymena thermophila* (~75% AT), an accomplishment that has not been reported in any other vector (E. Orias, personal communication). Importantly, this library was not stable in a similar BAC vector containing the *lac*Z-based blue/white screening (Figure 4).


pCC1BAC vector (8 kb)

Lucigen's pSMART VC vector (7 kb)

7 out of 18 clones deleted

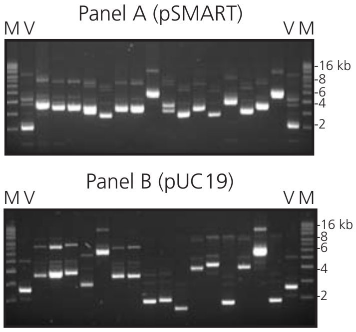
100% Stability

Figure 4. Genomic DNA from *Tetrahymena thermophila* (75% AT) was sheared, size-selected to 10-20 kb, and cloned into pSMART VC or a leading BAC/large insert vector. Uncut mini-prep DNA shows only large plasmids in the pSMART VC library, but many small, deleted clones in the conventional BAC vector.

Conclusions

Although pUC18 and its close relatives stably maintain most inserts, their limitations are becoming increasingly important with cloning of targets that may be unstable (e.g., AT-rich, highly repetitive, trace amounts, etc.). Numerous types of inserts are rendered "unclonable" by the high level of transcription and translation caused by the blue/white colony screen. In addition, transcription proceeding through inserts, or initiated within them, may lead to instability of the plasmid or failure of the blue/white screen. The pSMART vectors, designed specifically to circumvent many of the limitations of

continued, next page


Panel A (pSMART)

Panel B (pUC19)

Figure 2. Increased stability of AT-rich genomic clones in pSMART. A cosmid containing genomic DNA from *P. carinii* was hydrodynamically sheared, end repaired (DNATerminator® Kit; Lucigen) and shotgun-cloned into pSMART-HCKan or pUC19 (Lucigen). Plasmid DNA was purified from transformants by a standard alkaline lysis protocol. Uncut DNA from each clone was analyzed by gel electrophoresis (approximately 200 ng/lane, 1% agarose, TAE buffer). (Upper panel) Plasmids from randomly picked pSMART transformants were all within the expected size range, demonstrating high stability. (lower panel) Over 25% of the pUC19 transformants were unstable, yielding plasmids smaller than the parent vector. M, supercoiled plasmid ladder (Invitrogen); V, empty vector control.

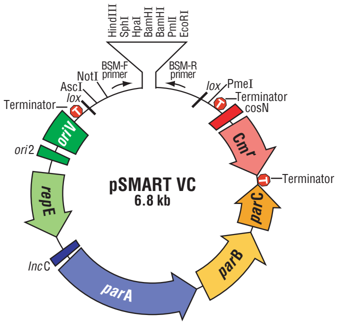


Figure 3. Diagram of the CopyRight™ Vector. The pSMART VC vector (transcription-free) is illustrated, showing positions of terminators (T), chloramphenicol-resistance gene (Cm^r), *cos* site, partition genes (*par*A, *par* B, *par*C), single copy origin (*ori*2, *inc*C, *rep*E), inducible origin (*ori*V), cloning site, and primer binding sites.

pUC plasmids, completely eliminate the blue/white screening and also incorporate transcriptional terminators, lower copy numbers, and resistance to antibiotics other than ampicillin. Such dedicated cloning vectors substantially reduced the bias against cloning a variety of difficult targets, allowing stable maintenance of strong promoter fragments or toxic coding sequences that are refractory to cloning in pUC-type vectors. Cloning into transcription-free vectors rather than pUC derivatives appears to substantially decrease the bias in shotgun libraries, facilitating more rapid assembly of genomic sequences.

References

1. Mead, D.A., and Kemper, B. 1988. Chimeric single-stranded DNA phage-plasmid cloning vectors, in Vectors: A Survey of Molecular Cloning Vectors and Their Uses. Rodriguez, R.L., and Denhardt, D.T., eds. Butterworth. Boston, 85-102.
2. Mardis, E., McPherson, J., Martienssen, R., et al., 2002. What is finished, and why does it matter. *Genome Res* 12: 669-671.
3. Vieira, J., and Messing, J. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* 19: 259-268.
4. Yount, B., Denison, M.R., Weiss, S.R., and Baric, R.S. 2002. Systematic assembly of a full length infectious cDNA of mouse hepatitis virus strain A59. *J Virol* 76: 12335-43.
5. Bowater, R.P., Jaworski, A., Larson, J.E., et al., 1997. Transcription increases the deletion frequency of long CTG·CAG triplet repeats from plasmids in *Escherichia coli. Nucleic Acids Res* 25: 2861-2868.
6. Wild, J., Hradencna, Z., and Szybalski, W. 2002. Conditionally amplifiable BACs: Switching from single-copy to high-copy vectors and genomic clones. Genome Res 12:1434-1444. ■

PERFECT CLONING!

*

First Time…

Every Time!

Only with

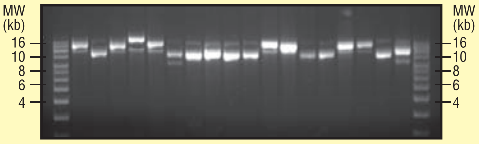
CLONESMART®

DNA Cloning Kits!

Get great results every time with *any* DNA.

- **Large Inserts**
- **Toxic genes**
- **AT- or GC-rich sequences**
- **Shotgun libraries**
- **cDNA**
- **PCR fragments**
- **Repeats & hairpins**
- **Trace amounts**

CloneSmart kits give you unbeatable results! Even the toughest DNA inserts are stable in pSMART® vectors (see below).



AT-rich *Bacillus cereus* large insert library in pSMART LCKan vector.

For more information, visit www.lucigen.com or call toll-free 1-888-575-9695.

* No humans were cloned in the production of this advertisement.

Perfect results with any DNA!

CloneSmart® Blunt Cloning Kits offer unprecedented power and ease for all cloning situations, from routine subcloning to making unbiased libraries from "unclonable" DNA.

More accurate results

The unique pSMART® transcription-free vectors (U.S. Pat. 6,709,861) in CloneSmart Kits eliminate the cloning bias inherent in other vectors. Even the most difficult genes or sequences are stable. With pSMART vectors, you don't worry about lost clones or altered sequences.

More recombinants with less work

pSMART vectors yield >99% recombinants. No blue/white colony screening is needed! With conventional vectors, up to 40% of the clones with inserts are discarded because they produce a blue color.

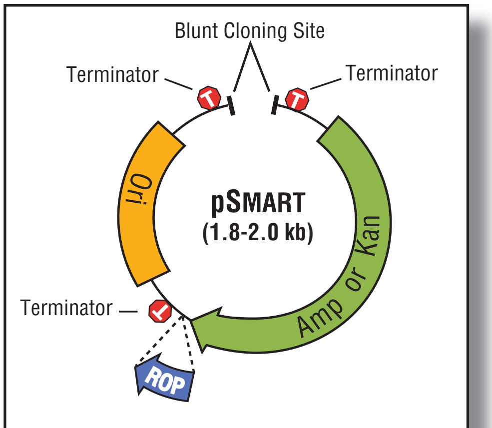
Ideal for all cloning applications

Difficult inserts. pSMART vectors maintain unstable or toxic fragments.

Small to large inserts. Clone DNAs from less than 100 bp up to 10-15 kb.

High fidelity shotgun libraries. CloneSmart Kits give much more comprehensive libraries and minimize sequence gaps.

Easy subcloning and mutagenesis. The small size of pSMART vectors allows simple and efficient subcloning, mutagenesis, and transposon insertion.



Transcription terminators prevent transcription into or out of the insert.

Maximum convenience

The power and simplicity of the CloneSmart system makes it ideal for all your cloning needs.

- Pre-processed vectors eliminate the need for digestion, gel purification, and dephosphorylation.
- High efficiency competent cells guarantee large, complex libraries.
- Cloning reagents are pre-mixed — just add your DNA.
- Ultra low background allows cloning of <10 ng for individual fragments, <500 ng for libraries.
- No post-ligation cleanup needed prior to transformation.
- pSMART vectors contain either ampicillin or kanamycin resistance, and high- or low-copy origins of replication.
- Compatible with all *E. coli* competent cells.

Lucigen®

The Molecular Cloning Company™

Protein Expression

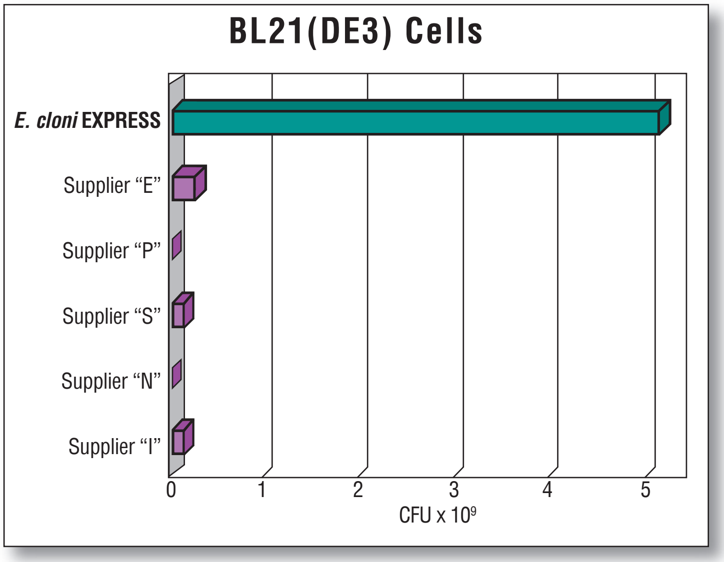
The Best Expression Yet!

Introducing *E. cloni*® EXPRESS Competent Cells

Lucigen's new *E. cloni* EXPRESS BL21(DE3) Competent Cells give you protein expression results faster, saving time, money and work. With a transformation efficiency 25- to 1,000-fold higher than similar cells from any other supplier, *E. cloni* EXPRESS allows you to clone and express in the same cell, without transferring plasmids from a cloning strain into an expression strain. Cloning and expression of single genes or entire libraries has never been this easy.

- Clone & express in the same cell
- Unparalleled transformation efficiency ($\geq 5 \times 10^9$ cfu/µg DNA)
- Available in BL21(DE3), BL21(DE3)pLysS, and BL21(DE3)pLysE versions. The pLys cells optimize expression of toxic proteins
- Conveniently packaged in 2 reactions per tube
- Includes Recovery Media and Positive Control DNA
- Also available: EXPRESS Chemically Competent Cells

E. cloni EXPRESS products are sold under a license issued to Lucigen Corporation by Brookhaven Science Associates, LLC (BSA) for research use by noncommercial entities. Commercial entities require a license from BSA.



To find out more about *E. cloni* EXPRESS and our other ultra efficient competent cells, call Lucigen at **1-888-575-9695** or find us on the web at **www.lucigen.com.**

Lucigen®

The Molecular Cloning Company™