Journal of Experimental Microbiology and Immunology (JEMI) Vol. 6:35-38

Polymerase chain reaction amplification of *lacZ* to engineer a gene sequence that can be used to modify the size of pUC19

CORINNE KRENTZ
Department of Microbiology and Immunology, UBC

In an attempt to increase the size of pUC19, the *lacZ* gene (3kb) was targeted for amplification from *Escherichia coli* C29 genomic DNA by PCR. Gel electrophoresis showed that a 3kb band was absent, and major products about 700bp and 1500bp were present. To determine what had been amplified, PCR products were cloned into a vector prior to sequencing. None of the clones contained the 1500bp product. Sequence analysis showed that the 700bp fragment is the *hisG* gene. Possibly due to the large size of the template and poor primer design, the PCR primers annealed to a different location and the desired construct was not obtained.

Evidence exists that co-transfection of pUC19 and pBR322 in *Escherichia coli* DH5α results in exclusion of pBR322 (W. Ramey, personal communication). Small plasmids like pUC19 (2686bp) can be replicated quickly and are produced in large quantities, whereas the replication of larger plasmids require more time and resources. Replication of large plasmids can cause stress on the cell, thus fewer copies of plasmid are made. The large size of pBR322 (4361bp) may explain the selection against its replication in cells co-transfected with pUC19 and pBR322. Using an engineered pUC19 that is the same size as or larger than pBR322 will enable us to conduct experiments to test whether plasmid size is a factor in the exclusion of pBR322 observed in *E.coli* DH5α.

The size of pUC19 will be increased to 5755bp with the addition of *lacZ* gene, which encodes β-galactosidase. The *lacZ* gene was selected because it is a large gene whose product can be easily detected by a colormetric methods including blue-white colony screening and it is known to exist in *E.coli* so it should not be harmful to the cell.

The *lacZ* gene was amplified by PCR from *E.coli* C29. The forward primer sequence was altered to create a *Bam*HI site. A convenient *Hin*dIII site already existed downstream from the *lacZ* gene. The unique *Bam*HI and *Hin*dIII sites that exist in the multiple cloning site of pUC19 will be utilized for insertion of the PCR product (Fig. 1).

MATERIALS AND METHODS

Primer Design. *E.coli* CFT073 (GenBenk accession number NC_004431) genome sequence was obtained from the National Center for Biotechnology Information (www.pubmed.com). Primers were designed with PrimerDesigner software using *E.coli* CFT073 genomic sequence. The forward primer sequence was 5' TTCACACAGGATCCAGCATG 3'. It had one base pair mismatch to create a *Bam*HI restriction site, one base pair deletion to keep the coding sequence in the correct reading frame, and the T_m is 62°C. The everse primer sequence was 5' GCGCTGTGTGCCTCGATTGT 3'. All twenty reverse primer bases are complementary to the coding strand and the T_m is 77°C.

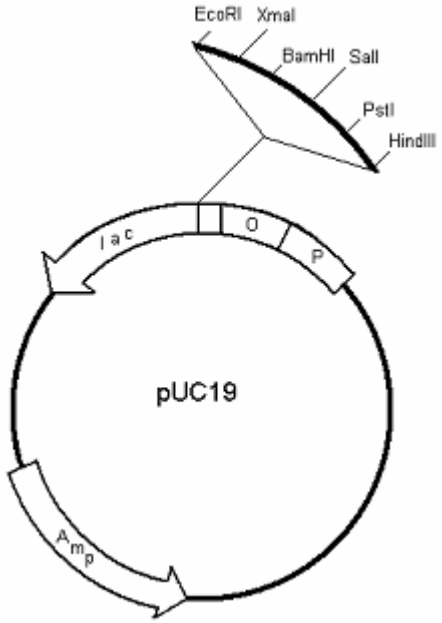

FIG. 1 pUC19 MCS. pUC19 and the PCR fragment are both cut with *Hin*dIII and *Bam*HI. A *Bam*HI sticky end will be generated at the start of the β-galactosidase gene and a *Hin*dIII sticky end at the end of the gene. The compatible sticky ends will join and the β-galactosidase gene will be inserted in the reverse orientation.

PCR. Genomic DNA was obtained from *E.coli* C29 (Hfr PO 12.20' CW tonA22, phoR19 (constit) ompF627 (T2R) fadL701 (T2R) relA1 pit-10 spoT1 rrnB-2 mcrB1) using phenol chloroform extraction technique (*4*). *E.coli* C29 genomic DNA was used as the template because it has a complete *lacZ* gene. Sixteen PCR reactions were run, varying template and $MgCl_2$ concentration. The template solution was diluted 1:1000, 1:100, 1:10, and undiluted. The $MgCl_2$ concentrations used were 1mM, 2mM, 3mM, and 4mM. Each reaction used 1.25 units of *Taq* DNA polymerase, 5uL 10X *Taq* buffer, 0.4uM of each forward and reverse primer, 1uL dNTP, and distilled water to bring the final volume to 50uL. The cycling conditions were initial denaturation 94°C for 3min; denature 94°C for

1min; anneal 45°C 1.5 min; extend 50°C 10min for 40 cycles; final extension 50°C for 20min. Gradient PCR was also used in an attempt to amplify the *lacZ* fragment in the temperature range of 42.5-67.5°C. The desired fragment was not obtained.

Cloning the PCR Product. Products amplified at 47.6°C and 65.1°C were ligated into a PCR-TOPO® vector according to the Invitrogen Life Science TOPO TA Cloning® Catalogue (*3*). TOPO Cloning® utilizes the property that *Taq* –amplified PCR products have a 3'-A overhang at each end (*3*). The 4uL PCR product, 1uL salt solution, and 1uL PCR-TOPO® vector were reacted for 40min. Sets of duplicate samples of the One Shot® Competent *E.coli* were heat shocked for 45 sec and for 2 min (*3*). The cells were plated on LB agar containing tryptone, yeast extract, and sodium chloride. The LB was supplemented with X-gal and 50μg/mL ampicillin. Light blue and white colonies were labeled 'a' to 'l' and grown in LB broth supplemented with 50μg/mL ampicillin. The PCR-TOPO® containing the PCR fragment insert were isolated from each culture using Concert High Purity Plasmid Purification System Mini-prep kit (*3*). All DNA samples had A_{260}/A_{280} above 1.8 and sufficient concentration for restriction enzyme digest.

DNA Digest. The purified plasmids were incubated with *Eco*R1 in REact3 buffer at 37°C for 18h to remove the PCR product from the plasmid. The digest products were separated on a 1% agarose gel in 1X TAE buffer.

DNA Sequencing. Samples from the TOPO clones E, F, and G were quantified and sent to the Nucleic Acid Protein Service Unit (NAPS) at the University of British Columbia for DNA sequencing analysis. Applied Biosystems PRISM 377 automated sequencers were used in combination with other Applied Biosystems equipment and chemistry to sequence the DNA.

Sequence Analysis. ClustalX version 1.83 was used to obtain a multiple sequence alignment of sequences E, F, and G. BLAST (www.ncbi.nlm.nih.gov/BLAST/) was used to identify the region of the *E.coli* genome that was amplified.

RESULTS

Amplification of the *lacZ* gene was attempted from *E.coli* C29 genomic DNA. The annealing temperature, template concentration, and $MgCl_2$ concentration were varied in an attempt to create optimal conditions for primer binding to the *lacZ* target. However, analysis by gel electrophoresis indicates that the expected fragment was not obtained (Fig. 2). The *lacZ* gene is about 3kb and the largest product that is seen on the gel is about 700bp.

Template concentration appeared to have no effect on the quantity of the desired 3kb PCR product. Higher $MgCl_2$ concentrations appeared to increase yields of the 700bp PCR product. Low concentrations of $MgCl_2$ appear to be insufficient to generate large amounts of PCR product. Gradient PCR showed that lower annealing temperatures also appear to amplify a 700bp fragment and higher annealing temperatures amplify a 1500bp product as well as a 700bp fragment (data not shown). In order to determine whether these smaller products were related to the *lacZ* sequence, the PCR products were cloned into pCR2.1 plasmids. *Eco*RI restriction digest of pCR2.1 containing a PCR product showed that some of the plasmid samples contained a 700bp insert (Fig. 3), but the 1500 bp fragment was not successfully cloned into the TOPO vector.

Three of the inserts in the clones with the 700bp insert were sequenced. A multiple sequence alignment of the TOPO clones E, F, and G was performed to identify PCR product, primer, and vector regions (Fig. 4). Surprisingly, clone G also contained a 700bp insert. The alignment shows indels in F that are not present in E and G, and indels in both E and G that are not present in F. Sequence E and G aligned nearly perfectly; however, it appears that F is in fact the same PCR product as E and G. All three sequences show a PCR product of 770bp, and sequence F is simply the complementary strand in the opposite orientation as E and G. The high degree of similarity between the three sequences supports the hypothesis that E, F, and G are the same PCR product of 770bp. A BLAST search of the PCR product sequence revealed that the amplified fragment corresponds to the middle section of the *his*G gene in *E.coli*. The sequence data also shows that only the reverse primer annealed to the *his*G template.

DISCUSSION

Primers were designed using *E.coli* CFT073, whose genome has been completely sequenced. Since the *E.coli* C29 genome has not been sequenced, the assumption was made that the C29 sequence would be conserved in the primer region. If the nucleotide sequence of C29 differs in the primer region, that would explain why the primers did not bind to the target region.

Primers are a critical element of PCR, and how well primers are designed will determine the success of the PCR. In future experiments, the T_m of primer pairs should not differ by more than 5°C (the primers used here differ by 15°C) and annealing temperature should be set 3-5°C below the T_m (*4*). Primer length can be increased to 25 nucleotides to increase specificity, and the ends of the primers should not have more than three G or C nucleotides, as this may contribute to nonspecific binding. Higher amounts of template have also been shown to increase nonspecific amplification (*2*).

A BLAST search of the *E.coli* genome was performed to identify other potential primer binding regions. No significant matches were found for either the forward or reverse primer. Since the 1500bp fragment was not successfully cloned and sequenced, we cannot definitively state that the primers did not anneal to the target region. It is possible that the *lacZ* gene was partially amplified due to secondary structure of template DNA. In future experiments, the 1500bp PCR product could be cut out of the gel and purified prior to cloning to improve the chances of cloning that PCR product.

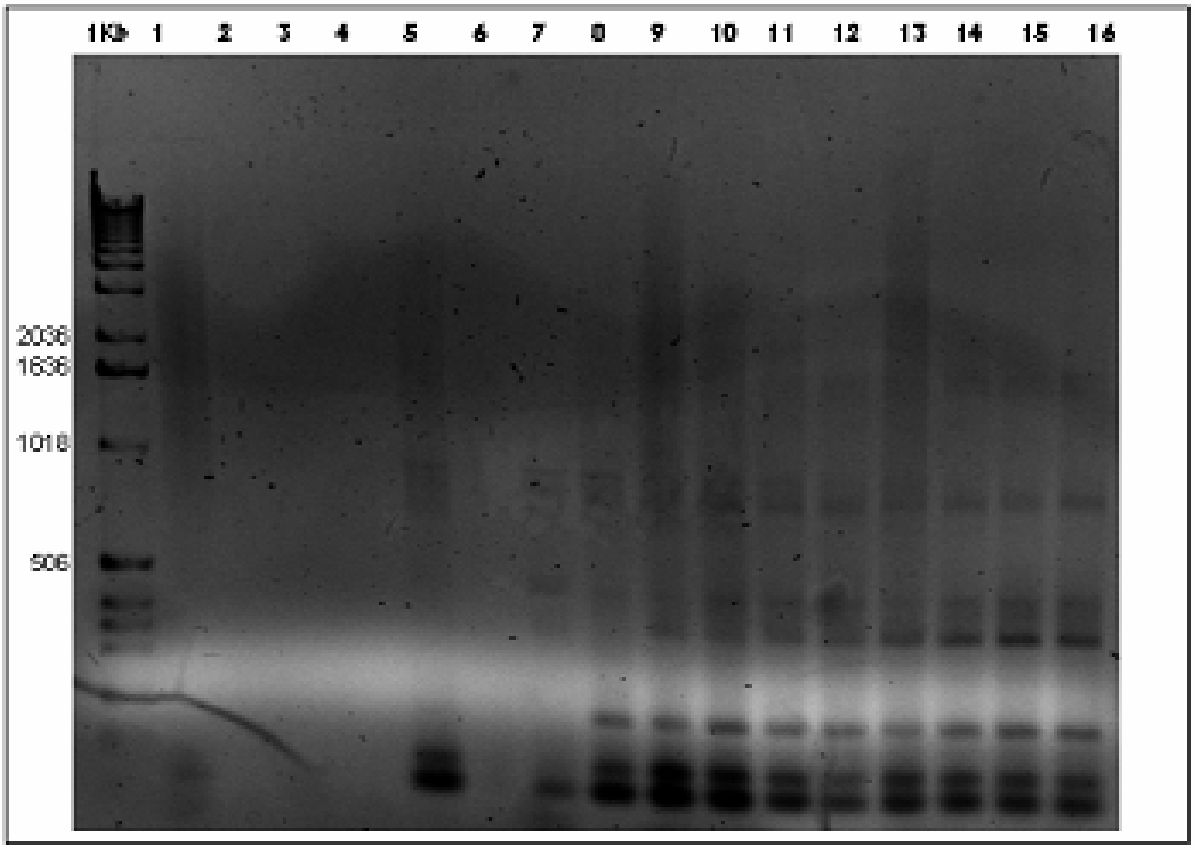

FIG. 2 Effects of magnesium chloride concentration on the PCR amplification of the *lac*Z sequence from *E.coli* C29 genomic DNA. Lanes 1-4 contain 1mM Mg^{2+}; lanes 5-8 contain 2mM Mg^{2+}; lanes 9-12 contain 3mM Mg^{2+}; lanes 13-16 contain 4mM Mg^{2+}.

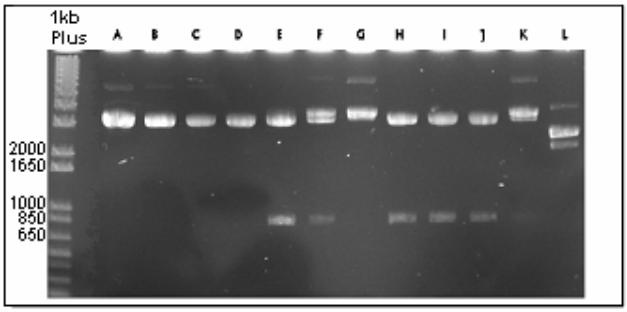

FIG. 3 Restriction endonuclease digests of positive TOPO clones of the PCR enriched products to identify clones with specific inserts. All samples have been cut with *Eco*R1. TOPO isolates E,F, H, I, and J appear to contain a 700bp insert.

An alternative to amplifying *lac*Z from the *E.coli* genome is to use a shorter template so there would be less extraneous genes that could potentially be amplified. In future experiments, the pβgal from Clontech or the low copy number F^{1}lac plasmid could be used as template (*1*).

When digested with *Eco*RI, the 770bp insert was removed from the vector in clones E and F but no such band was seen for clone G. Clone G was not expected to contain the same insert as E and F; however, the sequence results for clones E, F, and G showed that all three samples contained the 770bp PCR product, which is part of the *his*G gene. It is possible that the supercoiled nature of clone G DNA is responsible for the incomplete digestion. Only 2.5units of *Eco*RI per μg of DNA was used, and 5-10units/μg of DNA is recommended for supercoiled samples (*2*). Perhaps more enzyme was needed in order to achieve complete digestion.

Due to time constraints, experimental exclusion studies of pBR322 and modified pUC19 were not executed. The expected construct, which was intended to be cloned into pUC19, was not obtained by PCR.

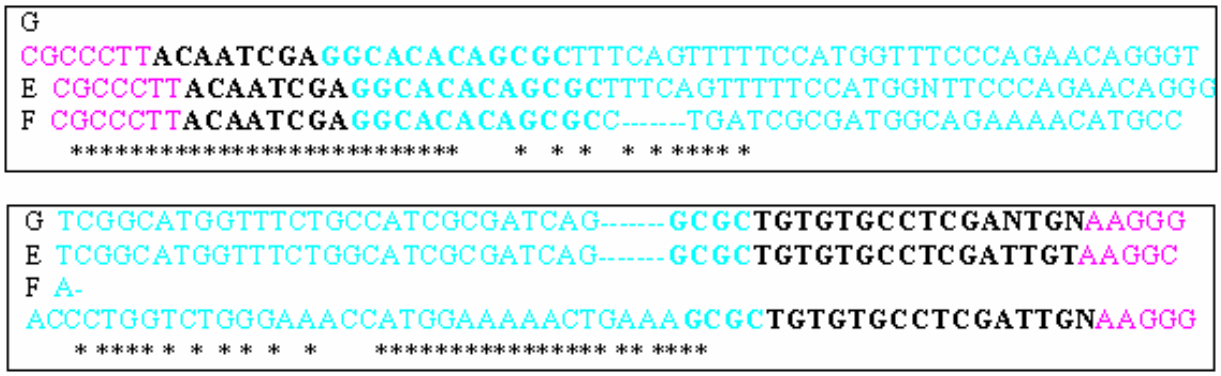

FIG. 4 Multiple sequence alignment for the TOPO clones E, F, and G shown in Figure 3. (Pink) pCR2.1 plasmid; (Green) PCR product; (Black) primer. The bold text shows where the reverse primer bound.

ACKNOWLEDGEMENTS

I would like to thank Jen Sibley and the other staff for their assistance, and Dr. Ramey for providing direction for this project. I would also like to thank the staff at Nucleic Acid Protein Service Unit (NAPS) for providing timely and high quality DNA sequence results.

REFERENCES

1. **BD Biosciences.** Clontech pβgal-Control. Retrieved January 5, 2004 from: http://www.bdbiosciences.com/clontech/techinfo/vectors/vectorsA-B/pbetagal-Control.shtml
2. **Fermentas Life Sciences.** Protocol for PCR and Restriction Endonucleases. Retrieved January 6, 2005 from: http://www.fermentas.com
3. **Invitrogen. TOPO® Cloning Technology.** Retrieved January 4, 2005 from: http://www.invitrogen.com/content/sfs/brochures/710_021849%20_B_TOPOCloning_bro.pdf
4. **Sambrook, J., Russell, D.** 2001. Molecular cloning: a laboratory manual, 3rd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.