Rolling Circle Mutagenesis of GST-mCherry to Understand Mutation, Gene Expression, and Regulation

Jessica Cole¹, Amanda Ferguson², Verónica A. Segarra³, and Susan Walsh²*
¹Department of Biology, Portland State University, Portland, OR 97207-0751, ²Department of Biology, Rollins College, Winter Park, FL 32789, ³Department of Biology, High Point University, High Point, NC 27268

Undergraduates are often familiar with textbook examples of human mutations that affect coding regions and the subsequent disorders, but they may struggle with understanding the implications of mutations in the regulatory regions of genes. We have designed a laboratory sequence that will allow students to explore the effect random mutagenesis can have on protein function, expression, and ultimately phenotype. Students design and perform a safe and time-efficient random mutagenesis experiment using error-prone rolling circular amplification of a plasmid expressing the inducible fusion protein glutathione S-transferase (GST)-mCherry. Mutagenized and wild-type control plasmid DNA, respectively, are then purified and transformed into bacteria to assess phenotypic changes. While bacteria transformed with the wild-type control should be pink, some bacterial colonies transformed with mutagenized plasmids will exhibit a different color. Students attempt to identify their mutations by isolating plasmid from these mutant colonies, sequencing, and comparing their mutant sequence to the wild-type sequence. Additionally, students evaluate the potential effects of mutations on protein production by inducing GST-mCherry expression in cultures, generating cell lysates, and analyzing them using SDS-PAGE. Students who have a phenotypic difference but do not obtain a coding region mutation will be able to think critically about plasmid structure and regulation outside of the gene sequence. Students who do not obtain bacterial transformants have the chance to contemplate how mutation of antibiotic resistance genes or replication origins may have contributed to their results. Overall, this series of laboratories exposes students to basic genetic techniques and helps them conceptualize mutation beyond coding regions.

INTRODUCTION

Background

One of the most important qualities of genetic material is that it replicates faithfully. DNA polymerase mediates this high fidelity. By distinguishing the appropriate deoxyribonucleotide triphosphates (dNTPs) from the surrounding environment, polymerases elongate complementary daughter strands of DNA. Base-pair size and shape define geometric selection, which pairs a specific dNTP to the appropriate base on the template strand. Quantitative analyses on polymerase fidelity have indicated that the enzymes perform with significant efficiency; for every million nucleotides synthesized, one error may occur (1). Additionally, DNA polymerases can often replace incorrect nucleotides and exchange them to attain the correct base pairing through proofreading (2). Magnesium ions (Mg²⁺) are a key cofactor for DNA polymerase activity and fidelity (3). While high fidelity polymerase activity is imperative to the natural process of replication, methods to reduce that fidelity have been employed in experiments to induce point mutations and directed evolution (4).

Random DNA mutagenesis is often used to explore the altered function of a specific protein. There are a variety of methods that can be used for achieving random mutagenesis of plasmid DNA, including chemical mutagens, mutator strains, and PCR (5). Chemical mutagens such as ethyl methanesulfonate (EMS) or nitrous acid are effective to modify DNA, but they are toxic and can be dangerous in the classroom (6). Another method uses mutator gene strains, such as E. coli mutD5, which operate by inducing a defect in DNA polymerase proofreading; however, this technique is difficult to employ because of the inherent genetic instability of the bacterial strains (7). In contrast, error-prone PCR, or ‘sloppy PCR,’ is an in vitro method that reduces DNA polymerase fidelity, generating a mutated gene product quickly and easily without toxic chemicals (4). One of the more efficient and accessible mutation strategies is error-prone rolling circle amplification (RCA) of plasmid DNA (8, 9). Like PCR, RCA amplifies circular DNA by using error-prone DNA polymerase and rolling circle amplification. RCA is a powerful tool for generating random mutations in both DNA and RNA templates (10, 11). RCA can be used for gene mutagenesis to generate random mutations for directed evolution (12, 13) or random mutagenesis to explore the potential effects of mutations on protein function (14).

©2017 Author(s). Published by the American Society for Microbiology. This is an Open Access article distributed under the terms of the Creative Commons Attribution-Noncommercial-NoDerivatives 4.0 International license (https://creativecommons.org/licenses/by-nc-nd/4.0/ and https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode), which grants the public the nonexclusive right to copy, distribute, or display the published work.
DNA, yet it does not require specific primers, and the reaction can occur at a single, maintained temperature (4). In vivo, RCA is utilized by some viruses and the F plasmid of E. coli (10). The process involves nicking the DNA and replicating around the circle, displacing the broken strand and using the unbroken strand as a template. Displaced single-strand DNA re-circularizes and can be used as a template to generate a new double-strand DNA circle. When manganese chloride or calcium chloride is added to the isothermal rolling circle reaction, DNA polymerase fidelity decreases (3-4, 11–13). The mutated, amplified, purified, and concentrated DNA product can be transformed into E. coli, giving rise to many clonal colonies, each containing a randomly mutagenized plasmid. DNA sequencing can then be used to analyze the plasmids for mutations. In contrast to other published procedures using fluorescent proteins (14), our technique is more affordable and open-ended since it employs a forward-genetics instead of reverse-genetics approach.

As a template for mutagenesis, we used the plasmid mCherry/pGEXKG. mCherry is a monomeric fluorescent protein (FP) derived from the naturally occurring FP, DsRed, through directed evolution. As its name suggests, mCherry fluoresces red (excitation = 587 nm; emission = 610 nm), is photostable, and matures quickly (15). As a nontoxic monomer, it is ideal for generating fusion proteins within a living cell. For example, bacteria that express high levels of a glutathione S-transferase (GST)-mCherry fusion protein appear pink even under normal lighting conditions. Like green fluorescent protein (GFP), the structure of mCherry is a β-barrel with a chromophore in the middle (16). Key amino acids for generating the acylimine chromophore include residues 66 to 68 (MYG), and mutation of these can result in changes to the color of the protein (15, 17). Other important amino acids include those that interact with the chromophore, such as K70, L83, and E215 (17). Certainly, any significant change to the protein structure or chromophore formation could result in loss of the red color.

In this study, error-prone RCA was used to rapidly and randomly mutate a pGEX plasmid containing the GST-mCherry fusion gene under the control of the isopropyl β-D-thiogalactopyranoside (IPTG)-inducible Ptac promoter. We added divalent cations (manganese or calcium chloride solutions) to reduce the fidelity of the φ29 DNA polymerase. The RCA product was used for direct transformation of E. coli. In comparison with a control lacking added divalent cations, phenotypic effects of the mutated plasmid were apparent, as some colonies were white or darker pink. Differences in protein expression were explored using IPTG induction. Basically, since the Ptac promoter is a fusion of the trp and lac operon promoters (18), this plasmid allows for the analysis of the function of these operons in vivo. The Ptac promoter is repressed by the lac repressor and derepressed by IPTG, so students can explore the concepts of transcriptional regulation on protein expression. Finally, sequencing the GST-mCherry genes validated mutations within the coding region. Other changes in phenotype were likely due to mutations in promoters. Overall, this technique allows for rapid mutation and screening of mutants.

Intended audience/prerequisite knowledge

This laboratory series is intended for undergraduate students in a Genetics course. We have used this in a 300-level course for sophomore, junior, and senior biology, biochemistry, and marine biology majors. These students have already completed a full year of introductory biology and introductory chemistry. Ideally, this exercise dovetails with concepts including DNA replication and PCR, the effect of mutation on phenotype, and prokaryotic gene expression, including the lac operon. In our courses, we have found that this experiment is best utilized after classwork covering the central dogma and during discussions of prokaryotic gene regulation. In our courses, the next topic is DNA replication and mutation, so this aligns with the last week of the experiment, which involves a sequence analysis. This laboratory exercise spans these topics and therefore begins during week four of a fifteen-week semester. Students will be challenged to design original experimental conditions, using appropriate control groups, to mutate a bacterial plasmid and identify the location of the mutation site. Thus, by the start of the lab, they should be familiar with the concepts of gene cloning, such as plasmids, in-frame, rolling circular amplification, IPTG, and fusion proteins. Prior experience with pipetting and sterile technique is encouraged.

Learning time

To complete the laboratory in its entirety will take five laboratory periods of approximately three hours, with some preparation the night before or the day after the lab (Appendices 1 and 4). We have also performed the laboratory without the protein induction (week 4, Appendix 1) if time is limited. Students are instructed to read primary literature papers pertaining to rolling circle amplification and mutagenesis (4, 11) before coming to the first laboratory. Students should look at the pGEX plasmid map and identify where the mCherry gene was cloned in relation to the GST coding sequence, as well as the function of the following parts within the plasmid: Amp’, pBR322 ori, and Ptac. Because the experiment entails examining an easy-to-see phenotypic change in the transformed bacteria, students are encouraged to preview the amino acid sequences of some fluorescent proteins to compare with mCherry. This inquiry-based learning will help the students construct a hypothesis about how the phenotype may change with possible mutations. This preliminary work can be assigned in the form of a worksheet before the lab begins (Appendix 1).

Learning objectives

The learning objectives are outlined in Table 1.
TABLE 1.
Learning objectives and assessment.

<table>
<thead>
<tr>
<th>Learning Objective</th>
<th>Successful Completion of Objective</th>
<th>Means of Assessment</th>
<th>Percent of Students Achieving LO$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design a controlled experiment based on primary literature</td>
<td>Negative control (water only) employed and only one variable tested at a time (either same concentration of different cations or different concentrations of the same cations)</td>
<td>Laboratory discussion before initiating the experiment</td>
<td>100$^b$</td>
</tr>
<tr>
<td>Generate mutations in a bacterial plasmid</td>
<td>Successful transformation with different colored colonies on plates$^c$</td>
<td>Laboratory data</td>
<td>91.7%</td>
</tr>
<tr>
<td>Classify mutations by phenotype (changes in protein expression) and genotype</td>
<td>Accurate and complete description of mechanisms to generate a mutant phenotype (both coding region and regulatory region mutations)</td>
<td>Laboratory paper (results and discussion)</td>
<td>41.7% of students identified a mutation in the coding region.</td>
</tr>
<tr>
<td>Analyze sequencing data using basic bioinformatics tools</td>
<td>Interpretation of SDS-PAGE gel of protein induction (identifies WT control protein at 56kDa and compares mutant protein for both molecular weight and amount)</td>
<td>Laboratory paper (figure legend and results and discussion)</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Alignment of appropriate reading frames of the mutant proteins with the reference protein sequence</td>
<td>Laboratory paper (figure)</td>
<td>95.8%</td>
</tr>
<tr>
<td></td>
<td>Identification and classification of mutations, if present, in the coding region</td>
<td>Laboratory paper (figure legend and results and discussion)</td>
<td>92.8% accurately classified mutations</td>
</tr>
</tbody>
</table>

$^a$Data from fall 2015 and 2016.
$^b$Students are not allowed to proceed until this is achieved.
$^c$If this skill fails for one group, typically others in the class will have additional colonies to analyze.
$^d$Protein induction was not performed in fall 2015 due to timing.
$^e$Students who did not obtain coding region mutations were not counted.

PROCEDURE

Materials

A graphical outline of the procedure is shown in Figure 1. Special materials include illustra TempliPhi DNA Amplification Kit (GE Healthcare Life Sciences #25640010); plasmid mCherry/pGEXKG (available by request from the corresponding author); Econospin Minelute Columns (Epoch Life Science 3010-250); phi29 DNA polymerase (New England Biolabs M0269S); and primers (Table 2). Fluorescent images were captured using a green stereomicroscope fluorescence adapter (Nightsea SFA-LFS-GR) and a digital microscope imager (Celestron 44421). Standard materials include molecular biology equipment such as calcium chloride competent E. coli, calcium chloride and manganese chloride solutions, LB plates with ampicillin, SDS-PAGE equipment, Coomassie stain, and IPTG. Detailed instructions and concentrations follow here, in the student handout (Appendix 1), and in the instructional prep sheet (Appendix 4).

Student instructions

Rolling circle amplification, mutagenesis, purification, and transformation. An illustra TempliPhi DNA Amplification Kit (GE Healthcare Life Sciences #25640010) was used in the RCA. Plasmid mCherry/pGEXKG (10 ng) was combined with 5 μL TempliPhi sample buffer, denatured by heating at 95°C for 3 minutes, and cooled to 25°C. Depending on student experimental design, divalent cations or water (0.5 μL) were added for a final concentration of 0.25 mM to 1.5 mM. Samples were then diluted with 5 μL TempliPhi reaction buffer and 2 units of phi29 DNA polymerase (New England Biolabs M0269S). Reactions were incubated for 16 hours at 30°C and inactivated at 65°C for 10 minutes. Samples were stored at 4°C until the next laboratory period. The reaction was purified using Econospin Minelute Columns (Epoch Life Science 3010-250) according to manufacturer’s instructions, transformed into calcium chloride competent NEB5α E. coli cells (New England Biolabs; 14), and plated onto LB agar with 100 μg/mL ampicillin. Colonies were
grown overnight at 37ºC and stored at 4ºC until the next laboratory period.

**Sequence analysis.** Single colonies with disparate phenotypes were used to inoculate LB broth with 100 μg/mL ampicillin and grown overnight at 37ºC. Plasmids were purified using a homemade miniprep procedure (20), but a standard miniprep kit, such as those offered by Qiagen (27104) or New England Biolabs (T1010), could also be employed. Miniprepped plasmids were sent for sequencing (Eurofins MWG Operon) with various primers (Table 2). Chromatograms were checked for accuracy and translated using Four Peaks Software: http://nucleobytes.com/4peaks/. Sequences were compared with a known GST sequence (GenBank accession number: ACM86784.1; https://www.ncbi.nlm.nih.gov/protein/ACM86784.1) and the known mCherry protein sequence on Protein Data Bank (http://www.rcsb.org/pdb/explore/remediatedSequence.do?structureId=2H5Q) using Multalin software (http://multalin.toulouse.inra.fr/multalin/); note that the GST-mCherry fusion protein contains an extra three amino acids (PGS) between the C-terminus of GST and the N-terminus of mCherry. Mutations, if detected, were compared with the protein structure in the Protein Data Bank.

**Phenotype analysis.** Individual colonies on agar plates were imaged on a stereomicroscope using a green stereomicroscope fluorescence adapter (Nightsea SFA-LFS-GR) and a digital microscope imager (Celestron 44421). For protein analysis, 100 μl overnight culture of single colonies was used to inoculate 3 mL fresh LB with ampicillin for 90 to 120 minutes at 37ºC. Uninduced samples (100 μl) were removed and centrifuged for 5 minutes at 4000 g. The supernatant was discarded, and the pellet was resuspended in 40 μl of SDS-PAGE sample buffer and stored at -20ºC. IPTG was added to 1 mM to induce protein expression for 3 hours at 37ºC. Induced samples (50 μl) were removed and processed as above. All samples were heated to 95ºC for approximately 3 minutes before resolving by 12% SDS-PAGE. Gels were stained with Coomassie Blue.

**Faculty instructions**

Faculty should be familiar with techniques involving bacterial transformation, plasmid DNA isolation, bacterial protein expression, and SDS-PAGE. Detailed procedures and answer keys for the student handout are provided in Appendix 3, and an explicit list of materials is included in Appendix 4. In our hands, homemade calcium chloride competent NEB5α E. coli (19) yielded a significantly higher transformation efficiency than purchased calcium chloride competent E. coli. For the RCA reaction, we allowed the students to do the math to determine their treatments but, to simplify things, provided stock solutions at 11 mM (0.5 mM final), 22 mM (1 mM final), or 33 mM (1.5 mM final) so that students added a consistent but small volume (0.5 μl) to each reaction. Additional stock solutions were made as students varied their experiment. Faculty should also familiarize themselves with sequence analysis software or use their own resources. As all the software indicated here is free, we prefer these resources. In the interest of keeping costs low, each lab student pair or group only sequenced and analyzed two putative mutants.

Additional timing outside of the scheduled laboratory period for students and faculty to set up an experiment or collect data are indicated in the student lab handout (Appendix 1) and the prep sheet (Appendix 4). If timing allows, both week 3 and the induction and collection of samples for week 4 can be done at the same time so that only one overnight culture is required, leaving one week to do both a protein gel and the sequence analysis or two weeks to spread these exercises out. To combine weeks 3 and 4, the uninduced sample is collected at the start of lab, and the induced sample is collected at the end of the lab period. This option may be better if students have limited access to the lab or cannot work independently.
Suggestions for determining student learning

Learning objectives and means of assessment are described in Table 1. For us, students generate a complete laboratory paper detailing their experimental design, methods, results, and discussion (Table 1). They should describe how the mutations they identified likely affected protein structure, and if no mutations were found, what else might have happened to generate the phenotype. Students also include a data table of colony counts detailing the particular treatment (or not) and the phenotype of the colonies (pink, white, dark pink, etc.). In addition, class data of colony phenotype can be used to assess the effect of divalent cations on the fidelity of DNA polymerase (Fig. 2). Although we used a draft of a laboratory paper to assess achievement of the learning objectives, alternative methods may be employed, such as a poster presentation, figure submission, lab notebook grading, or others.

Sample data

DNA polymerase fidelity can be reduced in a RCA reaction by incorporating certain divalent cations (4, 11–13). In this experiment, we performed RCA on mCherry/pGEXKG samples incorporated with manganese or calcium cations and transformed these reactions into E. coli (Fig. 3). Phenotypic variation can be measured by examining the colonies using brightfield and fluorescence microscopy (Fig. 3). Changes in phenotype can also be assessed using protein expression and SDS-PAGE assays (Fig. 4). In addition, the sequence of the coding region can be used to identify mutations (Fig. 5) or, if none are found, to predict alternative ways that the phenotype may have changed (e.g., mutations in the promoter or operator). Students can also make predictions about mutations they will not see because they will inhibit bacterial growth, such as mutations to the antibiotic resistance gene or the origin of replication.

Safety issues

On the first day of lab, students are instructed about lab safety, compliant with the ASM Guidelines for Biosafety in Teaching Laboratories, specifically those for working with...
BSL-1 microorganisms, and must sign a lab safety agreement confirming that they will abide by these guidelines. Students use NEBα E. coli K12 (BSL-1; fhuA2 Δ(argF-lacZ) U169 phoA glnV44 1180Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17; New England Biolabs #C2987). This bacterial strain is classified as not dangerous and should be handled with good industrial hygiene and safety practice, according to the manufacturer’s website. LB containing ampicillin and other chemical reagents is used, and students should therefore wear lab coats, gloves, closed-toed shoes, and protective eyewear during the experiment. Coomassie blue stain and destain can cause irritation to the eyes, skin, and respiratory system if protective gear is not worn. Disposal of bacteria and reagents should be contained to appropriate biohazard bins within the laboratory. Bacterial solutions and plates should be autoclaved or bleached before disposal. Compliant with ASM guidelines, students should maintain a clean workspace and thoroughly wash their hands before exiting each laboratory session.

DISCUSSION

Field testing

This activity was implemented in the spring 2014, fall 2014, fall 2015, and fall 2016 upper level general Genetics courses. A total of 63 students experienced this inquiry-based laboratory activity.

Evidence of student learning

Laboratory data and papers from two classes of students, fall 2015 and fall 2016, were analyzed to see whether the learning objectives were met (Table 1). Two student-generated figures are included in Appendix 2 as examples. In the first figure, the student accurately described the mutation, and in the second figure, the student identified no mutations, extrapolating that mutations generating the phenotype likely occurred in the promoter. The figures and figure legends demonstrate that the students were capable of generating mutations, using bioinformatics analysis to translate, align, and understand their mutations, and correlating the mutations (or lack thereof) to phenotype. The students compared their sequences to those of a control sequence.

FIGURE 4. A colony producing no color also produces no protein after induction with IPTG. (A) Control, white, and dark pink cultures are shown after induction with 1 mM IPTG. (B) Protein samples from the cultures in A were resolved by SDS-PAGE and stained with Coomassie blue. U is uninduced before IPTG, and I is induced. IPTG = isopropyl beta-D-thiogalactopyranoside.

FIGURE 5. Cherry coding region sequences of nine white colonies from the same RCA show variation in mutations. Mutations are summarized as follows: 460: missense K173E; 468: missense G138D; 467: missense G25D; 459: frameshift at 139; 470: missense G57D and frameshift at 94; 464: missense I12F and frameshift at 118. Three sequences are not shown since they did not have mutations in the Cherry sequence; these data do not rule out mutations in the GST coding region or the promoter. RCA = rolling circle amplification. GST = glutathione S-transferase.
Possible modifications

The intent behind these experiments is to rapidly mutate a bacterial plasmid so that cells containing these mutant plasmids can be phenotypically identified and then analyzed. One extension, if equipment is available, is to examine bacteria by fluorescence microscopy (Fig. 3C). In addition, if funds are available, primers could be designed to sequence the entire plasmid to generate a mutation rate for each condition or to sequence and analyze additional plasmids per student or lab group. Furthermore, students could perform this test on additional FP-fusion proteins if available. We used GST-mCherry because it is brighter than GST-EGFP under normal lighting conditions, but other FPs may work as well. Finally, not all aspects of this procedure must be completed if timing is difficult. For example, students still will see a change in phenotype without doing the protein expression and SDS-PAGE activity.

SUPPLEMENTAL MATERIALS

Appendix 1: Student handout
Appendix 2: Student-written figures
Appendix 3: Answer key for student handout (Appendix 1)
Appendix 4: Reagent/equipment list and instructional prep sheets

ACKNOWLEDGMENTS

This experiment was conducted in spring 2014, fall 2014, fall 2015, and fall 2016 Genetics courses, and the authors appreciate the efforts of the students. Funding was supplied by the Rollins College Biology Department. The authors declare that there are no conflicts of interest.

REFERENCES