



## Development of a Rapid, Single-Cell Method of Recombinant Clone Screening Using Flow Cytometry

Saelin Bjornson\*, Brian Shim\*, Grace Kuo\*, Bruno Freitas\*

\*All authors contributed equally. Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada

**SUMMARY** Fluorescence activated cell sorting methods have been adapted to accelerate the isolation of recombinant bacterial clones in the past decade. Previous studies have established methods for sorting bacterial cells transformed with plasmids encoding for one fluorescent reporter protein, however, to our knowledge, no system has been reported using more than one fluorescent marker. In this proof-of-concept study, we construct pGSBB, a dual-reporter plasmid encoding both enhanced green fluorescent protein (*egfp*) and *mCherry* that would enable high-throughput quantification and cell-sorting of bacterial transformants using flow cytometry and fluorescence activated cell sorting. We showed that *E. coli* transformed with pGSBB variants can be visualized as discrete populations of cells expressing either GFP, mCherry, both reporters, or neither using flow cytometry. Future researchers may leverage this construct to expedite conventional cloning protocols that use phenotypic-based screens requiring overnight colony growth.

### INTRODUCTION

As the potential benefits of understanding and engineering bacterial systems are becoming increasingly evident with various medical, economical, and environmental implications, the demand to further elucidate the governing molecular and genetic mechanisms is growing at an exponential rate (1). Among all bacteria, *Escherichia coli* is the most well-characterized model microbe due to its experimental practicality and its relevance to industrial uses and human health. Historically, extensive research conducted on *E. coli* isolates has profoundly enhanced the development of genetic engineering techniques and technologies, including molecular cloning and recombinant DNA (2). Cloning refers to the creation of many identical DNA molecules, generally accomplished by *in vitro* insertion of DNA into a plasmid cloning vector backbone. This is done by restriction endonuclease digestion of constitutive fragments followed by ligation, and the resulting DNA is referred to as recombinant. These newly generated plasmids are subsequently transformed into *E. coli* for replication, which are plated on medium containing plasmid-selectable markers to select transformed cells for growth (3). As a conventional practice, *E. coli* bacterial colonies are cultured on LB agar plates with the assumption that a single-cell gives rise to a pure colony. Currently, streaking colonies remains the preferred tool for phenotypic selection of recombinant clones.

However, several studies have determined streaking isolated colonies can be insufficient for obtaining pure cultures. In one study investigating *Geobacter sulfurreducens*, Shrestha *et al.* elucidated that contamination by rare variants in a bacterial population continue to persist following repeated purification by streaking (4). Furthermore, vibrational spectroscopy has found localization-dependent difference in gene expression and lipid metabolism within

**Published Online:** 13 September 2019

**Citation:** Bjornson S, Shim B, Kuo G, Freitas B. 2019. Development of a Rapid, Single-Cell Method of Recombinant Clone Screening Using Flow Cytometry. UJEMI+ 5:1-12

**Editor:** Julia Huggins, University of British Columbia

**Copyright:** © 2019 Undergraduate Journal of Experimental Microbiology and Immunology.

All Rights Reserved.

Address correspondence to:  
<https://jemi.microbiology.ubc.ca/>

microcolonies, implying significant biological variance and heterogeneity within a colony (5). Heterogeneity aside, culturing *E. coli* colonies on agar plates is often a rate-limiting step in cloning protocols. Rudimentary verification methods include insertion into a read-out gene such as *lacZ* in the case of blue-white screening (6) or conditionally lethal genes in the case of positive selection vectors (7). More laborious diagnostics involve isolating plasmids before digesting and resolving them on agarose gels, performing colony PCR for detection of inserts, or directly sequencing plasmids (8). These methods require the growth of single cells to colonies, thereby slowing workflow.

Screening and selection for genotypes or phenotypes of interest have become advanced and high-throughput, specifically in the field of metagenomics. As the majority of bacteria are unculturable (9), to circumvent this limitation, metagenomic genetic or functional screens isolate DNA directly from the environmental samples, from which large DNA fragments are cloned into large-insert vectors, such as cosmids, fosmids or bacterial artificial chromosomes (BACs) (10). These are then propagated in a host bacterium and inspected. Creating large-insert libraries is a laborious process, involving careful titrating of bacteriophage and dilution-plating to obtain the desired number of clones per plate (11). Additionally, in the case of metagenomic functional screens, homogeneous colonies are crucial to accurately screening for a function of interest, as a mixed population can mask phenotypes. Even among cells containing the same insert, differences in vector copy-number can define a threshold for detection (11). Taken together, these factors highlight the need for a high-throughput method for selecting single bacterial cells on the basis of the vector or vector-insert they contain, thereby conserving time, preserving homogeneity and increasing sensitivity.

Single-cell techniques have advanced tremendously in the past several decades. This revolution, in part, was driven by the recognition of the genetic, physiological, biochemical and phenotypic heterogeneity present within clonal populations. Isolating individual cells for single-cell analyses enables the characterization of discrete and unique features that may potentially be overlooked in bulk analyses (11). One method that is especially user-friendly and accessible is flow cytometry (FCM). Flow cytometry is a method for quantifying the intensity of fluorescent molecules in individual cells. The flow cytometer uses lasers to excite fluorochrome conjugated antibodies or recombinantly expressed fluorescent proteins, such as green fluorescent protein (GFP) and mCherry. Once excited, the proteins will emit light at unique wavelengths which can then be detected by the flow cytometer. The flow cytometer is also able to characterize cells based on their size and granularity. The forward scatter (FSC) parameter separates cell populations by size, while the side scatter (SSC) parameter separates cell populations by their granularity. Further, the flow cytometer is a high-throughput method of single cell analysis, as the machine is capable of accurately detecting cell populations at up to 10,000 cells per second. Combining these characteristics with the cell sorting capabilities of a Fluorescence-Activated Cell Sorting (FACS) cells can be sorted and enriched from a heterogeneous population (11).

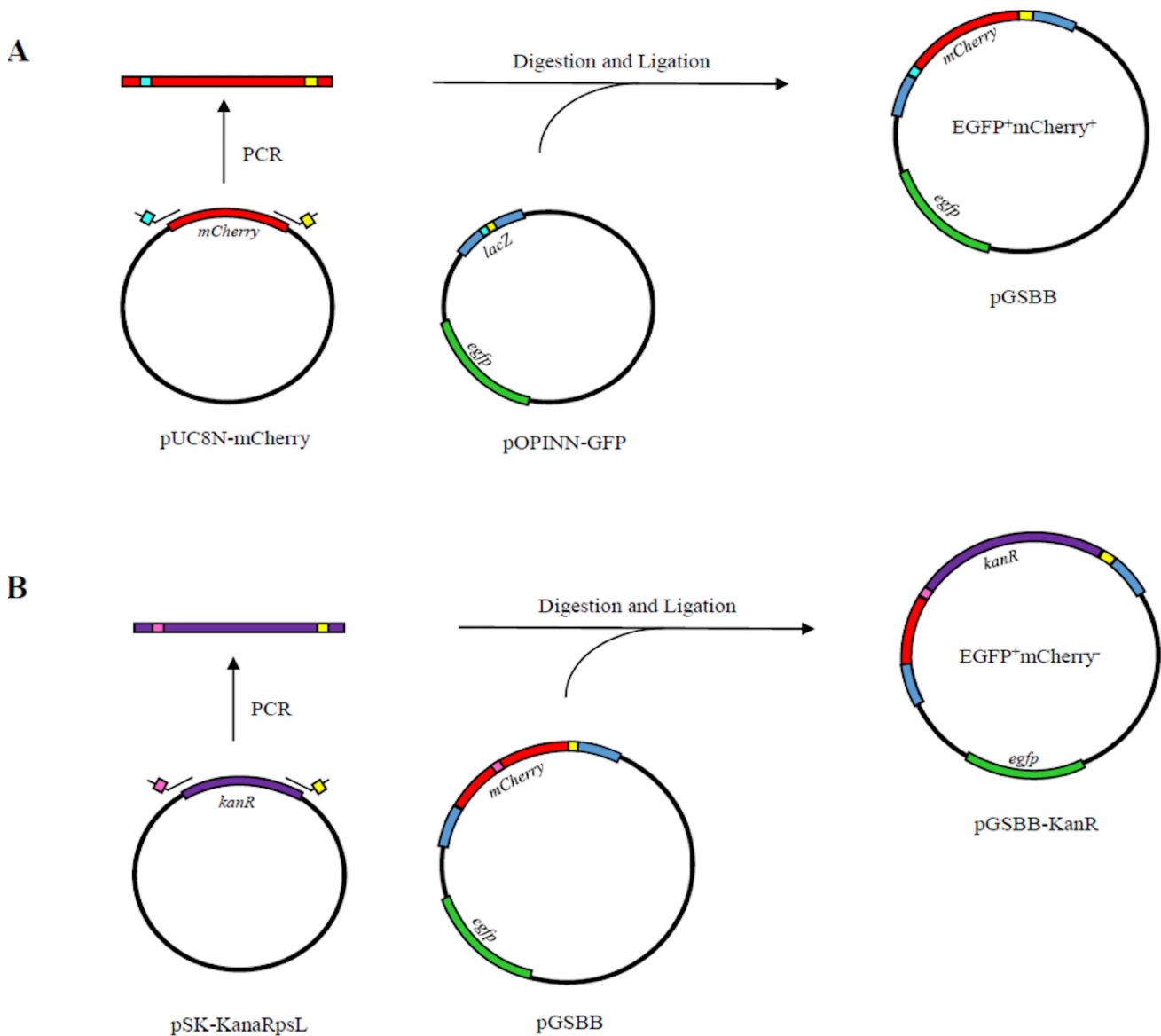
Enhanced Green Fluorescent Protein (EGFP) and mCherry proteins have been used for flow cytometric studies in Eukaryotic cells, with emission spectrums of each protein being detected from excitation by the 488 nm and 532 nm wavelength laser, respectively. The use of FCM for research in a bacterial setting has increased in recent years to conduct single-cell genomics (11, 12). Moreover, recombinant fluorescent-protein expressing *E. coli* have been used by Bennett *et al.* (2016) to investigate the kinetics of protein expression following transformation, concluding that flow cytometry could detect *gfp* expression 40 minutes after transformation and induction of the pGLO plasmid into *E. coli* BL21 DE3 (9, 13).

However, utilizing fluorescence in plasmid engineering remains largely unexplored, and designing and testing a reporter plasmid engineered to take advantage of FACS for isolation of pure transformants has not yet been attempted. Several single fluorescent systems have been reported, such as using protein-fusion constructs (14) or as a measure of promoter activation and strength (15, 16). Here, we modify this approach and construct a dual fluorescent plasmid that enables the use of flow cytometry and FACS to distinctly quantify and separate *E. coli* that are non-transformed, transformed with plasmids alone and transformed with a plasmid containing an insert.

## METHODS AND MATERIALS

**Strains and Media.** *E. coli* DH5a and *E. coli* BL21 DE3 strains were obtained from the Microbiology and Immunology Department at the University of British Columbia. Both strains were grown at 37°C in Luria-Bertani (LB) liquid broth or 1.5% LB agar plates at 37°C. To select for the bacterial population, 50ug/ml kanamycin or 100ug/ml ampicillin were used to supplement growth media. 0.1uM IPTG, and 20mg/ml X-gal were added for induction and blue-white screening purposes, respectively.

**Heat-shock transformations.** Preparation of chemically competent cells and heat-shock transformations were carried out as described by the *Chang et. al* protocol (2017). *E. coli* DH5a and BL21 DE3 were grown to an OD600 above 0.4, washed with 0.1M CaCl<sub>2</sub> and resuspended in 0.1M CaCl solution containing 15% glycerol, from which 50 uL aliquots were stored at -80°C for future use.



**FIG. 1 Schematic diagram of cloning workflow to construct pGSBB (EGFP-mCherry) and pGSBB-KanR (EGFP-mCherry) plasmids.** (A) PCR amplification of *mCherry* from pUC8N-mCherry with addition of KpnI (blue) and HindIII (yellow) recognition sites, located both in amplicons and within *lacZ* of pOPINN-GFP, followed by digestion and ligation of amplicon into pOPINN-GFP to generate pGSBB, expressing both mCherry and EGFP. (B) PCR amplification of *kanR* from pSK-KanaRpsL with addition of StuI (pink) and HindIII (yellow) recognition sites, located both in amplicons and within *mCherry* of pGSBB, followed by digestion and ligation of amplicon into pGSBB to generate pGSBB-KanR, expressing EGFP only.

For transformations, cells were thawed on ice and 1ng of DNA was added to each 50ul aliquot. After a 30-minute incubation on ice, cells were submersed in a 42°C heat bath for 30 seconds and then placed on ice for 2 minutes. The cells were then transferred to labeled glass tubes containing 1 mL of pre-warmed LB broth and were left shaking at 37°C for a 1-hour recovery period prior to plating. 100uL of transformed cells were spread-plated at 1:20 and 1:100 dilutions.

**Plasmid Preparation and Primer Design.** Plasmids used for construction of the dual reporter were pOPINN-GFP (Addgene #53541), pSK-KanaRpsL (Addgene #20871), and pUC8N-mCherry (obtained from Dr. Nomellini of the Microbiology and Immunology Department). Plasmids were received in *E. coli*, and plasmid extraction and purification were performed with EZNA® Plasmid Mini Kits (Omega Bio-Tec Cat. D6942).

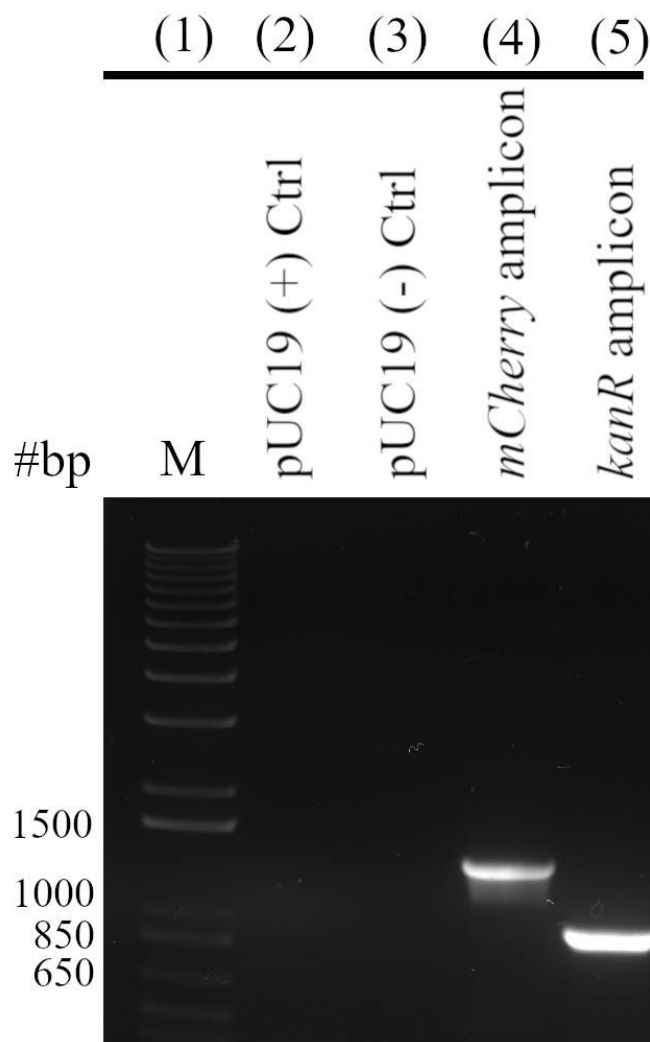
Primers for polymerase chain reactions (PCR) were designed using SnapGene Viewer 4.2.9 and the NCBI primer blast tool, and ordered from Integrated DNA Technologies (IDT). KpnI and HindIII restriction sites were added to the 5' ends of forward and reverse *mCherry* primers, while Stul and HindIII restriction sites were added to the 5' ends of forward and reverse *kanR* primers, respectively. As the KpnI restriction site in pOPINN-GFP is engineered for construction of EGFP-fusion proteins, a stop codon was added immediately following the KpnI in the *mCherry* forward primer to prevent run-through translation of EGFP into the insert. Primer sequences for PCR are found in Table S1.

**Amplification of *mCherry* and *kanR*.** PCR reactions were prepared in 50 uL dH<sub>2</sub>O with a final concentration of 1X Pfx amplification buffer, 50 mM MgSO<sub>4</sub>, and 1 Unit of Platinum® Pfx DNA polymerase (all provided as Invitrogen Cat. 11708039), 10mM dNTPs (Invitrogen Cat. R0193), with 0.3uM forward and reverse primers (diluted from 10uM stock) and 10 pg DNA template. Negative controls had dH<sub>2</sub>O in place of the template. Positive controls were carried out on isolated pUC19 with primer pairs targeting  $\beta$ -galactosidase, supplied by the Microbiology and Immunology Department.

The kanamycin resistance cassette (*kanR*) from pSK-KanaRpsL as well as the *mCherry* gene from pUC8N-mCherry, including its corresponding lac promoter and operator, were amplified by PCR. Reactions were carried out in a T100 thermocycler (Bio-Rad Laboratories) with the following conditions: 3 minutes at 95°C for initial denaturation, followed by 34 cycles of 30 seconds of denaturation at 95°C, 45 seconds of annealing at 56°C, and 1-minute extension at 72°C, with a final extension at 72°C for 5 minutes. Similar steps were programed for pUC19 positive controls, except the annealing temperature was set at 51°C due to the respective difference in primers melting temperature ( $T_m$ ).

**Agarose Gel Electrophoresis.** 1% agarose gels were made in TAE buffer with 1X SYBR™ Safe DNA gel stain (Invitrogen Cat. S33102). 10 uL of DNA (directly from PCR, ligation or digestion reactions) was mixed with 2 uL of 6X Orange DNA Loading Dye (Thermo Scientific Cat. R0611) and all 12uL was loaded into the wells. 100 bp and 1Kb Plus DNA ladder (Invitrogen cat. 10787018) were used for size estimation. Gels were run at 130V for 1 hour, after which they were visualized and imaged by Bio Rad Gel Doc™ System.

**Insertion of *mCherry* into pOPINN-GFP.** The *mCherry* amplicon and the purified pOPINN-GFP vector were double digested using KpnI (Thermo Scientific Cat. ER0521) and HindIII (NEB Cat. R0104S) restriction enzymes. Each 50 uL reaction contained 1X NEB CutSmart® Buffer (NEB Cat. B7204S), 1 Unit of each enzyme, 1 ug insert DNA and 1.6 ug vector DNA. Digestions took place at 37°C for 1 hour. Restriction enzymes were then heat-inactivated at 80°C for 20 minutes. Subsequently, pOPINN-GFP and *mCherry* digested products were ligated in a 3:1 amplicon-plasmid ratio in 20 uL reactions containing 1X T4 Ligase Reaction Buffer (Invitrogen #46300018), 1 Unit of T4 DNA Ligase (Invitrogen #15224017), 30 fmol of vector and 90 fmol of insert. Ligation reactions were incubated at room temperature for 2 hours. The ligated products were then transformed into 50uL of competent *E. coli* BL21 DE3, and plated on LB-Amp plates in the presence of IPTG and X-gal. The plates were incubated at 37°C for 16 hours. Colonies containing the dual reporter were indicated by both the lack of blue color from blue-white screening and the presence of



**FIG. 2 PCR amplification of *mCherry* from pUC8N-*mCherry* and *kanR* from pSK-KanaRpsL.** Expected and observed sizes are around 1300 bp for *mCherry* and 800 bp for *kanR*. Lane (1) Invitrogen 1kb+ Ladder; Lane (2) pUC19 positive control; Lane (3) pUC19 negative control; Lane (4) *mCherry* PCR amplicon; Lane (5) *kanR* PCR amplicon.

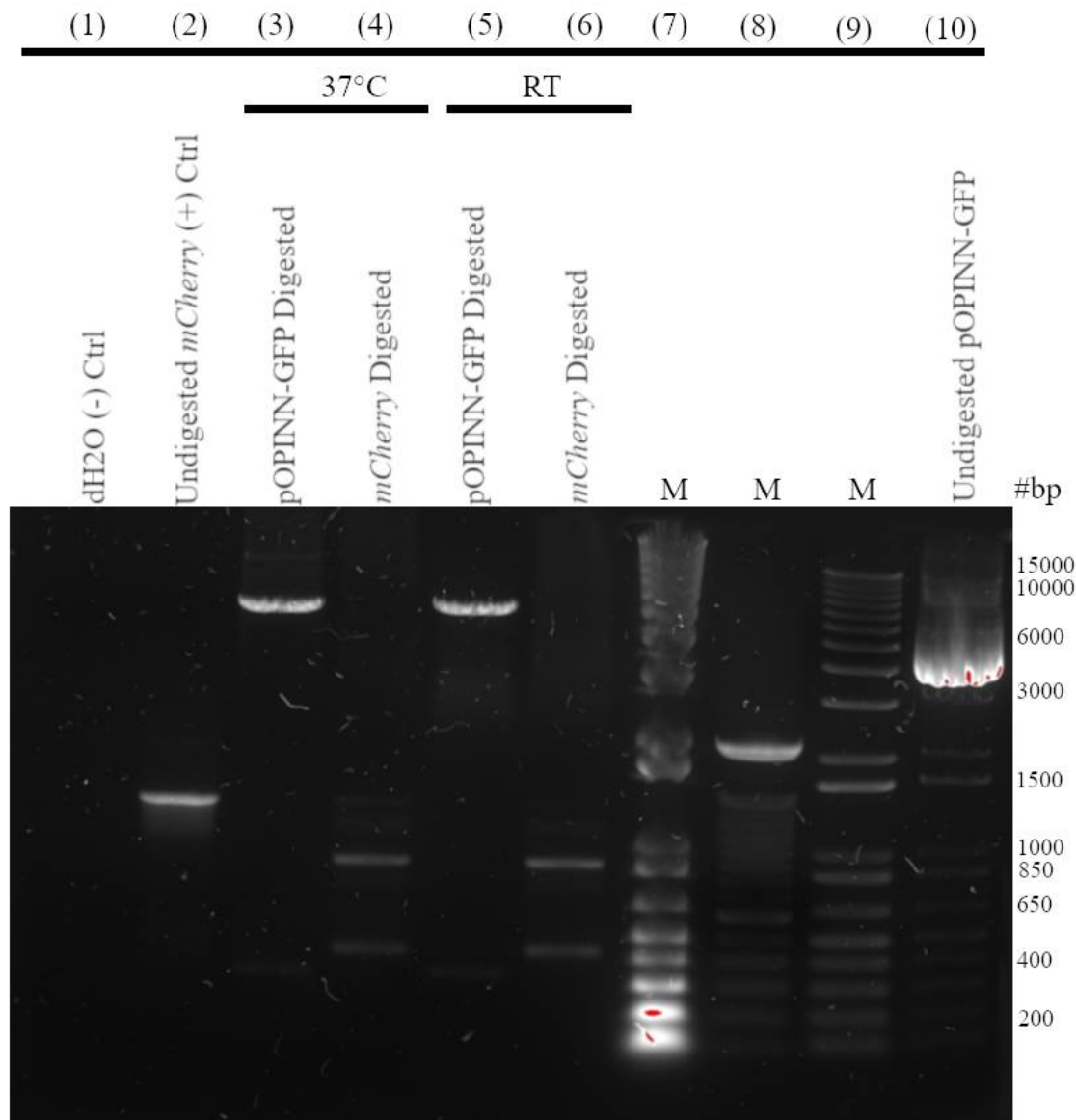
red color as the phenotypes of *mCherry*-expressing colonies. The dual reporter plasmid pOPINN-GFP-*mCherry* will be denoted as pGSBB for the subsequent analysis.

**Insertion of *kanR* into *mCherry* of pOPINN-GFP-*mCherry* (pGSBB).** The *kanR* amplicon and the purified pGSBB vector were double digested using *Stu*I (Thermo Scientific Cat. ER0421) and *Hind*III (NEB Cat. R0104S) restriction enzymes. Each 50  $\mu$ L reaction contained 1X CutSmart® Buffer (NEB Cat. B7204S), 1 Unit of each enzyme, 1  $\mu$ g insert DNA and 1.6  $\mu$ g vector DNA. Digestions took place at 37°C for 1 hour. Restriction enzymes were then heat-inactivated at 80°C for 20 minutes. pGSBB and *kanR* digested products were ligated in a 3:1 amplicon-plasmid ratio in 20  $\mu$ L reactions containing 1X T4 Ligase Reaction Buffer (Invitrogen #46300018), 1 Unit of T4 DNA Ligase (Invitrogen #15224017), 30 fmol of vector and 90 fmol of insert. Ligations were incubated for 2 hours at room temperature. The ligated products were then transformed into *E. coli* BL21 DE3 and plated on LB-Amp and LB-Amp-Kan plates in the presence of IPTG and X-gal at 37°C for 16 hours.

**Flow Cytometry Analysis.** *E. coli* BL21 DE3 cells with empty vector, pUC8N-*mCherry* only, pOPINN-GFP only, and pGSBB conditions were cultured overnight in liquid LB at 37°C, pelleted at 1200 relative centrifugal force (RCF) and resuspended in 1 mL of 1X TE buffer. The cells were then analyzed on the Attune™ NxT Acoustic Focusing Cytometer, blue/red/violet/yellow (Thermo Fisher Scientific, Waltham, MA, USA). Bacterial cells were gated using Forward Scatter (FSC-H) and Side Scatter (SSC-H). The voltages were set by running the empty vector control to establish the negative control voltage, then the pOPINN-

GFP and pUC8N-mCherry single positive controls to establish the EGFP and mCherry voltages. EGFP was excited by a 488 nm laser, and mCherry by a 586 nm laser, therefore no compensations were needed due to the minimal spectral overlap between the lasers. After running and recording all of the samples, the fcs files were exported and analyzed using FlowJO v10 Software.

FACS two-dimensional dot plots were used to analyze EGFP and mCherry fluorescence with respect to each other. By plotting EGFP fluorescence intensity (y-axis) against mCherry fluorescence (x-axis), the data makes it possible to differentiate between *E. coli* BL21 DE3 cells that express only one of the two fluorescent markers (top left quadrant for EGFP single positive; bottom right quadrant for mCherry single positive), those that express neither (bottom left quadrant), and those that express both (top right quadrant).



**FIG. 3 Visualization of pOPINN-GFP plasmid and mCherry amplicon using gel electrophoresis after double digestion with HindIII and KpnI.** Lane (1) negative control; lane (2) undigested mCherry positive control; lane (3) HindIII and KpnI digested pOPINN-GFP plasmid; lane (4) HindIII and KpnI digested *mCherry* amplicon; lanes (5) and (6) are the same as lanes (4) and (5), with the exception of a room temperature digestion period instead of the required 37°C; lane (7) Invitrogen 1kb+ Ladder; Lane (8) Invitrogen 100bp+ Ladder; Lane (9) Invitrogen 1kb+ Ladder; Lane (10) undigested pOPINN-GFP positive control.

### Sequencing of *mCherry* and *kanR* from pGSBB

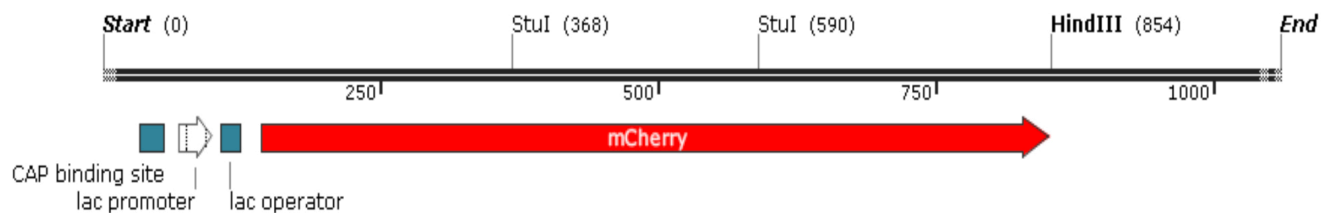
Plasmids from a pure colony of transformants resistant to ampicillin and kanamycin were extracted and quantified as previously described in the plasmid preparation section. Samples were prepared according to Genewiz sample submission guidelines by mixing 800 ng of plasmid DNA template with 25 pmol of sequencing primer. Two PCR tubes each containing the plasmid DNA with either *mCherry* or *kanR* forward PCR primers were submitted for sanger sequencing.

## RESULTS

**Cloning Workflow.** Figure 1 shows the cloning workflow to construct the dual reporter plasmid as well the steps required to knockout the *mCherry* region of this plasmid through insertion of *kanR*. The dual reporter was generated through the insertion of an *mCherry* amplicon into the *lacZ* region of pOPINN-GFP. To do this, *mCherry* of pUC8N-*mCherry* was PCR-amplified with incorporation of KpnI and HindIII recognition sites by designing them into the 5' ends of PCR primers, which are also located in the pOPINN-GFP *lacZ* region, as shown in Fig 1A. Once inserted, the new dual reporter construct was named pGSBB and contains fully functional *mCherry* and *egfp* genes, and was validated as EGFP<sup>+</sup>*mCherry*<sup>+</sup> using flow cytometry. Using the StuI and HindIII restriction sites located in *mCherry*, a *kanR* gene cassette was cloned into the *mCherry* region. *KanR* was PCR-amplified from pSK-KanaRpsL with incorporation of StuI and HindIII sites, as shown in Fig 1B. Insertion of *kanR* will disrupt *mCherry* function, and following transformation into *E. coli* result in cells that are EGFP<sup>+</sup>*mCherry*<sup>-</sup> when seen on the flow cytometer. These cells would then be sorted and grown on LB-Kanamycin plates, further validating that *kanR* ligated correctly and is expressed. Then, by comparing kanamycin resistance of FACS-sorted populations EGFP<sup>+</sup>*mCherry*<sup>-</sup>, the pGSBB dual reporter system will be confirmed as a viable tool for cloning

**Amplification of the *kanR* gene and *mCherry* gene and promoter region.** First, the *kanR* from the pSK-KanaRpsL plasmid and *mCherry* from the pUC8N-*mCherry* plasmid were PCR amplified, as shown in Figure 2. The sizes of the *mCherry* amplicon with its *lac* promoter and operator, approximately 1300 bp, and *kanR* is 800 bp, were as expected (Lanes 4 and 5). Lanes 2 and 3 are positive and negative technical controls of B-galactosidase amplification from the pUC19 plasmid, of which the expected band size is approximately 200 bp. It should be noted that negative (no-template) controls were not performed for *mCherry* or *kanR*, nevertheless because we observed bands of the expected size for the *mCherry* and *kanR* cassette we excised the bands from a gel and moved forward with cloning

**Cloning *mCherry* into the pOPINN-GFP vector.** To create the pGSBB construct, the *mCherry* gene and the pOPINN-GFP were double-digested by HindIII and KpnI restriction enzymes. Figure 3 shows the DNA gel electrophoresis performed post-digestion. Lanes 1 and 2 display a lack of band in the negative control, and the undigested *mCherry* positive control, respectively. The gel confirmed the proper digestion of both the plasmid (Lane 3) and amplicon (Lane 4), as two bright bands seen in Lane 3 match the expected band sizes of 6245bp and 400 bp that should be observed post digestion of pOPINN-GFP. These two bands correspond to the linearized plasmid (6245 bp) that remains after the removal of *lacZ* region (400 bp). In contrast to the digested pOPINN-GFP, the undigested plasmid DNA (Lane 10) has a lower band size indicating a faster migration rate, likely due to the supercoiled conformation. Two bright bands were also observed in the digested *mCherry* amplicon (Lane 4). The *mCherry* amplicon contains a HindIII restriction site just downstream of the gene. The two resulting bands from this digestion should appear at approximately 400 bp and 900 bp, as observed. After digesting both the insert and the vector, the *mCherry* amplicon was inserted into the pOPINN-GFP. The ligated products were then heat-shock transformed into *E. coli* BL21 competent cells. The presence of visibly white-red colonies after plating and incubating on LB-Ampicillin (100ug/mL) plates, containing IPTG and X-gal confirmed the successful ligation and proper orientation of the *mCherry* gene into the pOPINN-GFP

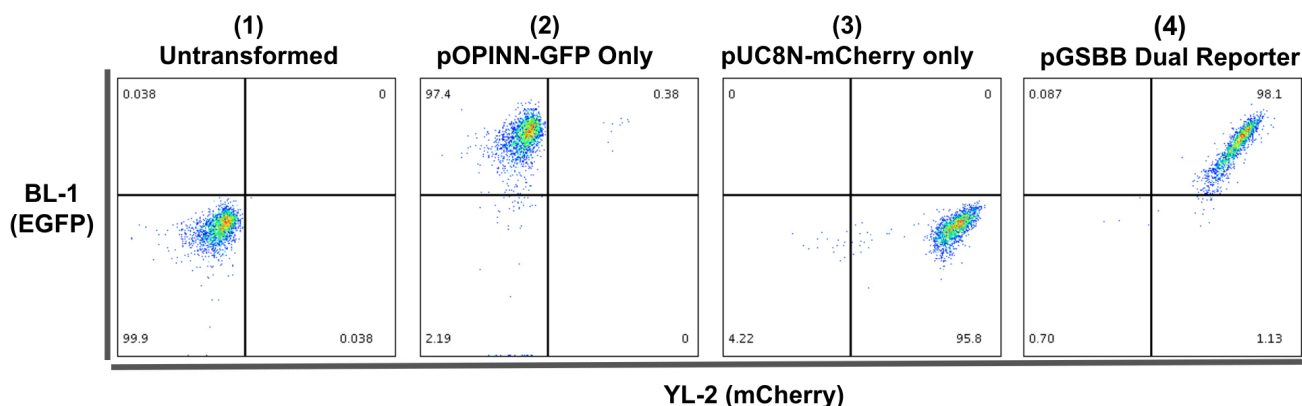


**FIG. 4 Sequencing results of *mCherry* gene using the PCR *mCherry* forward primer.** Confirmed the presence and location of the *StuI* and *HindIII* cut sites within/relative to the *mCherry* open reading frame.

plasmid. The visible expression of mCherry (Supplemental Figure 1.) in the presence of IPTG and X-gal suggests the *mCherry* gene is regulated by the upstream a 5' *lac* promoter and *lac* operator. The white colonies may be due to heterogenous expression dynamics.

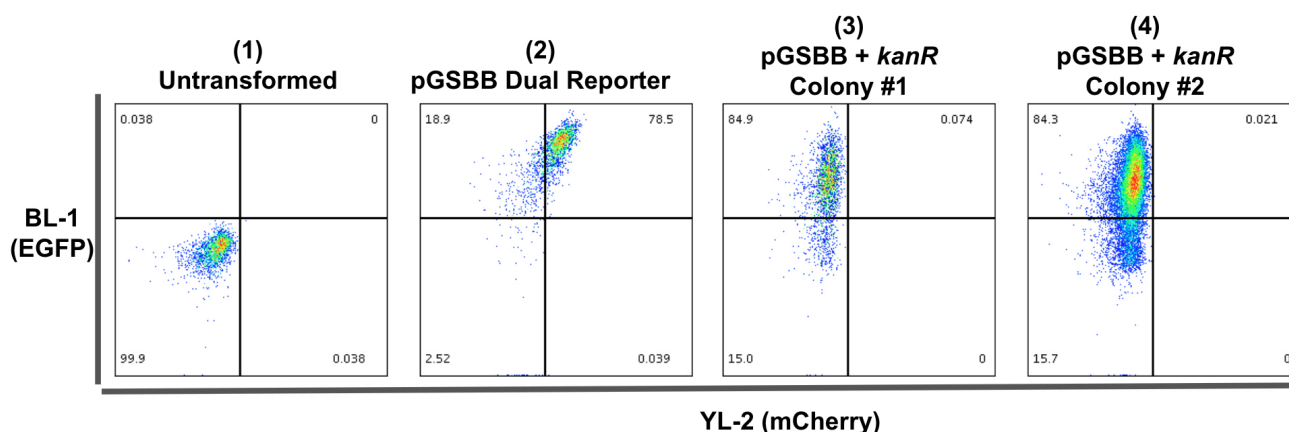
Sanger sequencing (Fig. 4) further confirmed the presence of the *mCherry* gene in the pOPINN-GFP, showing a full length *mCherry* gene as well as a 5' *lac* promoter and *lac* operator, two intragenic *StuI* cut sites, and a 3' *HindIII* cut site.

***E. coli* BL21 DE3 transformed with pGSBB are fluorescently distinguishable from single-fluorescent cells.** Conceptually, pGSBB was designed to differentiate non-recombinant from recombinant *E. coli* BL21 DE3 cells as EGFP<sup>+</sup>mCherry<sup>+</sup> and EGFP<sup>+</sup>mCherry<sup>-</sup>, as outlined in figure 1A and 1B, respectively. As a preliminary test, pUC8N-mCherry, pOPINN-GFP, and pGSBB colonies were grown on LB ampicillin plates, and cultured overnight in liquid LB at 37°C, and resuspended in TE buffer for subsequent flow cytometric analysis. Figure 5 shows the results of the test experiment. The EGFP and mCherry gates were set based on the localization of non-fluorescent, untransformed samples on the FCM dot-plots. Gate placements were further validated by the absence of single-fluorescent cells in the EGFP<sup>+</sup>mCherry<sup>+</sup> quadrant. The pGSBB clones were 98% EGFP<sup>+</sup>mCherry<sup>+</sup>, whereas the pOPINN-GFP clones were 97% EGFP<sup>+</sup>mCherry<sup>-</sup>. The small percentage of cells that are seen in the other quadrants likely are cells that have stuck to the walls of the flow cytometer from prior samples, thus resulting in false positives and negatives when a new sample is acquired. These errors can be mitigated by running 5% bleach between each sample, which was not done here. The experiment shows that flow cytometry can be used to differentiate single and double positive fluorescent BL21 DE3 *E. coli* cells transformed with variants of pGSBB.



**FIG. 5 Visualization of untransformed (1), and transformed *E. coli* BL21 DE3 with the pOPINN-GFP (2), and pUC8N-mCherry (3) and pGSBB dual reporter (4) plasmids using flow cytometry (left to right).** The y-axis represents EGFP fluorescence through the BL-1 channel, while the x-axis represents mCherry fluorescence through the YL-2 channel using the Attune™ NxT Acoustic Focusing Cytometer. Flow cytometry shows the expected outcomes for each condition: no fluorescence for the untransformed samples, EGFP only fluorescence for the pOPINN-GFP transformed samples, mCherry only fluorescence for the pUC8N-mCherry transformed samples, and double EGFP and mCherry fluorescence for the pGSBB transformed samples.

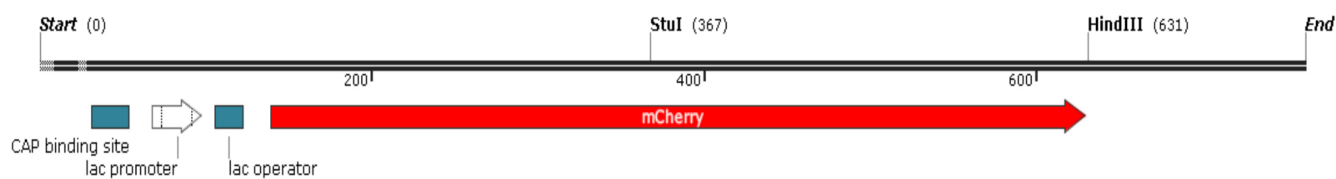




**FIG. 6 Visualization of untransformed (1), and transformed *E. coli* BL21 DE3 with the pGSBB Dual reporter (2), pGSBB+*kanR* plasmids isolated from Colony #1 (3) and Colony #2 (4) using flow cytometry (left to right).** The y-axis represents EGFP fluorescence through the BL-1 channel, while the x-axis represents mCherry fluorescence through the YL-2 channel using the Attune™ NxT Acoustic Focusing Cytometer. Flow cytometry shows the expected outcomes for each condition: no fluorescence for the untransformed samples, EGFP and mCherry fluorescence for the pGSBB transformed samples, and loss of mCherry fluorescence on two colonies isolated following ligation of *kanR* into *mCherry*.

***mCherry* gene knocked out following *kanR* and pGSBB ligation.** To test whether pGSBB indeed functions the way it was designed to, the pGSBB plasmid and the *kanR* amplicon were double-digested with *StuI* and *HindIII*, and the *kanR* was ligated into the open reading frame of *mCherry* in the plasmid. After ligation, the products were transformed into *E. coli* BL21 DE3. Half of the samples were plated on LB-Kanamycin plates, while the other half was inoculated in liquid LB-Amp for cytometric analysis. From flow cytometric analysis, the *mCherry* gene was shown to be knocked out after digestion of pGSBB and ligation of *kanR* and pGSBB (Fig. 6). Two colonies grown and picked from the LB-Amp plate displayed a cell population that is approximately 84% EGFP<sup>+</sup>mCherry<sup>-</sup>, and these cells are present as a distinct population when compared to the EGFP<sup>+</sup>mCherry<sup>+</sup> pGSBB dual reporter plasmid. The results show the expected absence of mCherry expression after ligation of *kanR* into *mCherry*. However, the samples grown on LB-Kanamycin plates to confirm the expression of Kanamycin resistance did not show growth of any observable colonies. It is possible that the kanamycin resistance gene disrupted *mCherry* expression as expected but was not expressed.

Since cells transformed with the ligation products *kanR* and pGSBB were shown to be EGFP<sup>+</sup> mCherry<sup>-</sup> but did not grow in the presence of kanamycin, we sequenced the plasmid to confirm that the *kanR* cassette had inserted into *mCherry* in pGSBB. Sequencing revealed that the two *StuI* sites within *mCherry* had ligated together, and no *kanR* sequence was detected (Fig. 7). With *StuI* being a blunt end restriction enzyme, the two *StuI* sites (Fig 4.) are compatible and thus made it possible for re-ligation to occur. Although the sequenced clone was not the target clone containing *kanamycin* insert, obtaining the *StuI* variant nevertheless fulfils our study objective of establishing a protocol for sorting *E. coli* transformants containing pGSBB variants using FACS. The flow cytometry demonstrated promising ability to visualize discrete populations of bacterial cells expressing either GFP, mCherry, both reporters, or neither.



**FIG. 7 Sequencing of pGSBB ligated products using *mCherry* forward primer.** Results showed ligation of the two original *StuI* sites in the pGSBB plasmid after double digest with *StuI* and *HindIII*.

## DISCUSSION

The objective of this study was to construct a novel dual-fluorescent plasmid, which we have named pGSBB, would enable rapid detection of transformants with the ability to distinguish between recombinant clones, non-recombinant clones and untransformed cells. Given the ability of a fluorescence activated cell sorter to simultaneously sort multiple populations, the pGSBB dual-reporter system would enable synchronous purification of recombinant clones from the overall population. This provides more dimensionality over the previously established single fluorescence systems (14).

Through our project, we have constructed pGSBB, a dual-reporter plasmid which encodes two distinct fluorescent proteins, *egfp* and *mCherry*. This dual-reporter system contains unique restriction enzyme cut sites both outside (KpnI and Acc65I for *egfp*; HindIII for *mCherry*) and within (PmlI for *egfp* and StuI for *mCherry*) each reporter protein, making it suitable for cloning applications. The use of mCherry and EGFP offers minimal spectral overlap as EGFP can be detected using the 488 nm wavelength channel, and mCherry can be detected using the 532 nm wavelength channel. Both channels are available in most commercially available flow cytometers and fluorescence activated cell sorters.

The compatibility of the flow cytometer and the dual reporter system was validated using single-fluorescent, non-fluorescent and pGSBB controls. The ability to knockout *mCherry* expression, resulting in a population shift from EGFP<sup>+</sup>mCherry<sup>+</sup> expressing cells to EGFP<sup>+</sup>mCherry<sup>-</sup> expressing cells was demonstrated (Fig. 6). This population is distinct from mCherry single positive, EGFP-mCherry double positive, and no fluorescence controls, which suggests these cells may be enriched for by using cell sorting. It should be noted that these different populations have not yet been analyzed as a mixed population; however, we have observed distinct populations when a ligation reaction was transformed into competent BL21 DE3 suggesting that this step is likely to yield acceptable results.

The pGSBB dual-reporter created in this study currently has some limitations that could be enhanced in future iterations of the pGSBB prototype. First, the pOPINN-GFP backbone that was used to construct pGSBB expresses EGFP under a T7 promoter. This restricts the use of this plasmid within cell lines that recombinantly express the bacteriophage T7 polymerase. Second, to ensure proper expression of an insert, cloning in a promoter with a gene is likely required. As the users would be limited to clone using exactly and only the two restriction sites offered by the pGSBB plasmid (StuI and HindIII), ideally, for ease of cloning, a multiple cloning site (MCS) containing unique restriction sites would be inserted within *mCherry* in a way that would not impact its fluorescence. If possible, this should be done close enough to the *mCherry lac* promoter to enable inducible expression of an insert. Additionally, the pGSBB plasmid may not be compatible for mammalian *in vivo* screens and clinical trials as exogenous mCherry and eGFP may induce an undesired immune response.

It is important to note that our method does not address expression of a gene once inserted, as its presence is implied indirectly by lack of expression of the fluorescence marker it inserts into. Our method is therefore limited to simply detecting the presence of insert within the recombinant plasmid when there are numerous critical factors affecting gene expression, protein synthesis and folding *in vivo* to result in a functional gene product (18). It should also be noted that the vector has the capacity to re-ligate and result in deletion of the *mCherry* gene and the generation of false positive colonies (Fig. 6 and 7) Therefore, we recommend phosphatase treatment of vectors during ligation to prevent this, which we were not able to undertake due to time constraints. In our case however, proper digestion should have resulted in some successful ligation, as HindIII creates sticky ends and StuI creates blunt ends. Because sticky ends are easier to ligate, *kanR* ligation should have out-competed the re-ligation of the two StuI sites. Digestion with HindIII may not have been successful, possibly due to the different buffer requirements of each enzyme. Our choice of the StuI restriction site in the *mCherry* gene for cloning may have been a poor one as KanR would be adjacent to the considerably large N-terminus of the mCherry protein. This may prevent folding of a functional KanR protein through steric hindrance. Nevertheless, the lack of successful insertion hindered our ability to test the sorting capabilities of FACS, which would ideally be done by comparing kanamycin resistance between sorted populations.

Flow cytometry is also an open system technique and hence is prone to contamination. We did not test contamination rates in samples that pass through the flow cytometer. More

complex methods for cell isolation, such as microfluidic cell sorting, do offer increased sterility, but are significantly costlier and requires special training (11). Additionally, potential morphological and genetic changes to the *E. coli* cells due to repeated laser exposure during enrichment process by FACS should also be taken into precaution by performing SEM imaging and genomic sequencing to ensure that the protocol does not induce mutations to the bacterial cells.

**Conclusions** Here we report a novel dual-reporter system containing two distinct fluorescent proteins, *egfp* and *mCherry*. This dual-reporter system contains restriction enzyme cut sites both outside and within each reporter protein, making it suitable for cloning applications. Moreover, flow cytometry has been shown to detect the emission of mCherry and/or EGFP in *E. coli* BL21 DE3 distinctively without having spectral overlap in the signals. EGFP can be detected using the 488 nm wavelength channel, and mCherry can be detected using the 532 nm wavelength channel. Both channels are available in a variety of flow cytometers and fluorescence activated cell sorters. Further, the flow cytometer is capable of detecting the presence of a reporter-gene knockout upon insertion of a gene of interest into the *mCherry* gene of the dual reporter plasmid. This population is distinct from the mCherry single positive, EGFP-mCherry double positive, and no fluorescence controls. However, *E. coli* BL21 DE3 transformed with this plasmid did not grow on LB-Kanamycin plates. Sanger sequencing demonstrated that the *kanR* cassette is absent in the plasmid, thus confirming that the issue is with the digestion of *kanR* during the cloning process of the gene. Nevertheless, pGSBB proves a likely candidate for enabling cell sorting from a heterogenous population of all three fluorescent phenotypes, but future studies will need to be done for validation.

**Future Directions** Based on our preliminary data, assays should be performed to determine a minimum-fluorescent-population limit of detection (sensitivity), as well as to assess the accuracy of flow cytometry in reporting population percentages of a heterogeneous sample (specificity). Bennett *et al* (2015) have shown that the expression of EGFP can be observed 40 minutes post induction of a pGLO plasmid, the kinetics and rate-limiting steps of this system can also be analyzed to determine an appropriate time frame for flow cytometry post-transformation, and to understand important biological variables and limits of our system when augmenting it for other uses. The enrichment percentage and sorting accuracy of the FACS machine can also be calculated by mixing different known ratios of the transformants containing the pGSBB, pOPINN-GFP and pUC8N-mCherry and comparing to population ratios after sorting.

Another application of the dual reporter system would be to determine the promoter strength in a plasmid via monitoring the fluorescence intensity in a time course experiment. The strength of *egfp*- and *mCherry*- associated promoters can be evaluated by RT-qPCR as an assessment of transcription rate whereas the rate of incorporation of radiolabeled amino acids into both proteins could be done to assess translation rate. Earliest observation of fluorescence (together with translation rates) would indicate rates of protein folding, and fluorescent signal strength can be quantified over time through time-course analysis experiments.

Once verified, our technique could significantly streamline a wide range of cloning techniques and applications. One such improvement would be in the preparation of metagenomic libraries. Fluorescent labeling of the commonly-used cloning fosmid pCC1FOS would enable the sorting of bacteria with linearized, insert-containing fosmids directly out of transduction media rather than plating on selection media for recombinant clones. Fosmid induction for high copy-number is often required for sufficient DNA yield or thorough screening of insert protein function, and enrichment of high copy-number bacteria by highly-expressing fluorescent markers through FACS would enhance these processes. For applications such as this one, it would also be imperative to assess the ability of FACS to sort not only plasmid-containing cells but cells with high copy-number, based on levels of fluorescence.

## ACKNOWLEDGEMENTS

This study was made possible by the funding from Microbiology and Immunology at the University of British Columbia. The authors would like to thank Dr. David Oliver and Mihai Cirstea for providing the support and guidance throughout the term. We would also like to acknowledge Dr. John Smit and Dr. John Nomellini for kindly providing the pUC8N-mCherry plasmid. We would also like to thank two anonymous reviewers for constructive feedback on this manuscript.

## CONTRIBUTIONS

BS and BF were primarily responsible for performing flow cytometry experiments. GK and SB were primarily responsible for conducting cloning experiments for generation of the pGSBB plasmid. All authors contributed equally both in the lab and during the writing period. The individual sections of the report were split equally among the authors: BS was in charge of the Abstract and Introduction, SB was in charge of the Materials and Methods, BF was in charge of the Results, and GK was in charge of the Discussion. The Figures, Tables, Conclusions and Future Directions were completed by each author.

## REFERENCES

1. Blount, Zachary D. 2015. The Unexhausted Potential of *E. coli*. *Elife*. 4:e05826
2. Lodish, H. F. 2016. Molecular cell biology Eighth Edition. New York: W.H. Freeman-Macmillan Learning.
3. Cohen, Stanley N., et al. 1973. Construction of Biologically Functional Bacterial Plasmids in Vitro. *Proc. Natl. Acad. Sci. U.S.A.*, 70:3240-3244.
4. Shrestha, PM, Nevin, KP, Shrestha, M, Lovley, DR. 2013. When is a microbial culture “pure”? Persistent cryptic contaminant escapes detection even with deep genome sequencing. *MBio*. 4:591.
5. Choo-Smith, L, Maquelin, K, Van Vreeswijk, T, Bruining, HA, Puppels, GJ, Thi, NN, Kirschner, C, Naumann, D, Ami, D, Villa, AM. 2001. Investigating microbial (micro) colony heterogeneity by vibrational spectroscopy. *Appl. Environ. Microbiol.* 67:1461-1469.
6. Lerner, CG, Inouye, M. 1990. Low copy number plasmids for regulated low-level expression of cloned genes in *Escherichia coli* with blue/white insert screening capability. *Nucleic Acids Res.* 18:4631.
7. Dower, WJ, Miller, JF, Ragsdale, CW. 1988. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.* 16:6127-6145.
8. Bennett, D, Chau, CH, Ma, R. 2015. Investigating Flow Cytometry as a Potential Method for Real-time Analysis of Gene Expression following *Escherichia coli* Transformation. *JEMI*. 19:
9. Stewart, EJ. 2012. Growing Unculturable Bacteria. *J. Bacteriol.* 194:4151-4160.
10. Lam, KN, Cheng, J, Engel, K, Neufeld, JD, Charles, TC. 2015. Current and future resources for functional metagenomics. *Frontiers in Microbiology*. 6:1196.
11. Blainey, PC. 2013. The future is now: single-cell genomics of bacteria and archaea. *FEMS Microbiol. Rev.* 37:407-427.
12. Yoon, HS, Price, DC, Stepanauskas, R, Rajah, VD, Sieracki, ME, Wilson, WH, Yang, EC, Duffy, S, Bhattacharya, D. 2011. Single-cell genomics reveals organismal interactions in uncultivated marine protists. *Science*. 332:714-717.
13. Abuaita, BH, Withey, JH. 2011. Genetic screening for bacterial mutants in liquid growth media by fluorescence-activated cell sorting. *J. Microbiol. Methods*. 84:109-113.
14. Ambriz-Aviña, V, Contreras-Garduño, JA, Pedraza-Reyes, M. 2014. Applications of flow cytometry to characterize bacterial physiological responses. *BioMed Research International*. 2014:.
15. Trip, H, van der Veek, Patricia J, Renniers, TC, Meima, R, Sagt, CM, Mohrmann, L, Kuipers, OP. 2011. A novel screening system for secretion of heterologous proteins in *Bacillus subtilis*. *Microbial Biotechnology*. 4:673-682.
16. Qin, JY, Zhang, L, Cliff, KL, Hulur, I, Xiang, AP, Ren, B, Lahn, BT. 2010. Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PLoS One*. 5:e10611.
17. Ferrari, BC, Oregaard, G, Sørensen, SJ. 2004. Recovery of GFP-labeled bacteria for culturing and molecular analysis after cell sorting using a benchtop flow cytometer. *Microb. Ecol.* 48:239-245.
18. Fakruddin, M. et al. 2013. Critical Factors Affecting the Success of Cloning, Expression, and Mass Production of Enzymes by Recombinant *E. coli*. *ISRN Biotechnology*. 2013:590587.