Undergraduate Journal of Experimental Microbiology and Immunology (UJEMI) Vol. 28:1-11

Functional characterization of IPTG and Salicylate inducible systems promotes utility of SIGEX Duo-directional Reporter Vectors in Metagenomic screening

Ryan Hong, Griff Wong, Jason Zhao, and Tian Hui Zhao

Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada

SUMMARY Precisely controlling and tuning inducible promoter responses to environmental stimuli is an important step in the engineering of synthetic gene circuits and biological systems for metagenomic applications. A duo-directional SIGEX plasmid using GFP and RFP reporters has been designed to address the unidirectional nature of biosensors. Studies have provided proof-of-concepts for IPTG and salicylate inducible pSPPH21-modified biosensors (pSPPH21-*lacI* and pSPPH21::*nahR/Psal*) by detection of GFP and RFP. Our incomplete understanding of SIGEX's sensitivity to different concentrations of inducers remains unknown. In this study, we explore the reporter expression dynamics of two SIGEX vectors controlled by different inducible promoters to investigate relationships between concentration and the effect on downstream gene reporter expression. We found that pSPPH21-*lacI* is responsive and hence sensitive across a vast range of IPTG concentrations ranging from 10 μM to 1 mM, with greatest inducible promoter responses seen between 8-10 hours. We also found that the pSPPH21::*nahR/Psal* plasmid is non-functional and is unable to be induced by salicylate. In sum, our work shows that SIGEX is sensitive across a wide range of inducible substrate concentrations. With a functional characterization of SIGEX established, further optimization steps can now be undertaken in preparation for high-throughput DNA promoter screening from an experimental metagenomic sample.

INTRODUCTION

Metagenomics has enabled an unprecedented growth in our understanding of the diversity of microbial life on Earth. This field has allowed for screening of genes-of-interest as well as confirming metabolic, physiological, and ecological functions of microbes without the isolation or cultivation of individual microorganisms. This has provided a greater degree of genomic information of microbial species compared to traditional, culture-based approaches (1-3). Substrate-induced gene expression (SIGEX) was first introduced and pioneered by Uchiyama and Watanabe as a method for isolating novel catabolic operons from metagenomic samples. SIGEX functions as a novel promoter trap method that harnesses a gene reporter system whose expression is induced by the presence of environmental stimuli, such as substrates or catabolites (1-2). Ultimately, SIGEX is a powerful method for not only metabolite detection and promoter discovery, but it also lends itself to high-throughput screening, especially when paired with other methods such as fluorescence-activated cell sorting (FACS) and other cell-sorting methods (3-4).

Previous studies have had great success in demonstrating that the duo-directional SIGEX vector is suitable for detecting gene expression activity induced by isopropyl-β-D-thiogalactopyranoside (IPTG) in *E. coli* using oppositely-oriented fluorescent reporter markers, GFP or RFP. However, these studies have only established a proof-of-concept to validate that the pSPPH21 vector can express both GFP and RFP upon induction with the allolactose analog, IPTG (5-7). These studies do not tell us: 1) how different concentrations of an inducible substrate and exposure time influences the expression or fluorescence of downstream reporter markers; 2) how microbial growth changes under different nutritional or stress environments (6-9). The *lac* operon system is very well-characterized, but little is known about the dynamic expression range of *lac*-inducible promoter systems and their effects on the cellular transcription levels of downstream reporter proteins, RFP and GFP, in a SIGEX (pSPPH21) context. The lack of functional characterization of these promoter-trap

Published Online: September 2023

Citation: Hong, Wong, Zhao, Zhao. 2023. Functional characterization of IPTG and Salicylate inducible systems promotes utility of SIGEX Duo-directional Reporter Vectors in Metagenomic screening. UJEMI 28:1-11

Editor: Lauren Pugsley and Shruti Sandilya, University of British Columbia

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

methods hinders our ability to predictably control transcriptional responses of SIGEX-based approaches.

SIGEX methods function largely on the assumption that inducible promoter systems can elicit precisely programmable responses. An important step towards engineering biological systems, such as SIGEX, is the ability to precisely tune promoter-inducible responses to environmental stimuli. Here, we functionally characterize the expression dynamics of two SIGEX vectors controlled by two different inducible promoter architectures, the *lac* operon and *nahR/PsaI*, induced by IPTG and salicylate respectively. We first validate the construct of lac operon containing pSPPH21 and pSPPH21::*nahR/Psal* using whole-plasmid sequencing. Next, we functionally characterize these two promoter architectures using a series of microplate reader experiments consisting of: 1) characterizing optical density (OD) measurements of microbial growth to investigate growth under different inducible substrate concentrations; and 2) describing the magnitude of gene reporter expression of a cloned *E. coli* host by creating dose-response relationships, whereby dose is the amount of inducible substrate (IPTG or salicylate) and response is the fluorescence intensity (GFP and RFP) after a certain exposure time. By providing a deep, functional characterization of these two inducible promoter systems, we hope this will aid in gaining a better functional understanding of the SIGEX scheme in terms of its efficacy and utility as a promoter trap for function-based metagenomic screening tools.

METHODS AND MATERIALS

Bacterial Strains and Plasmids

pSPPH21 strains stored in *E. coli* DH5α and were obtained on a starter plate from UBC MICB 471 laboratory stocks. *E. coli* DH5a containing the pSPPH21-lacI vector were obtained from Gawol et al. *E. coli* DH5a containing pSPPH21::nahR/PsaI from c#1, c#6.1, and c#13 were obtained from Frese et al. in glycerol stock.

Luria Bertani (LB) and chloramphenicol (CmR) media preparation. LB broth was prepared using 5g of tryptone, 2.5 of yeast extract, and 5g of NaCl dissolved in 500 ml distilled water, autoclaved and then stored at room temperature for future use. To create agar plates, 20 g of agar per liter of media was added to the LB mixture prior to autoclaving. LB agar plates with 20 ng/uL CmR were prepared using a 34 mg/mL chloramphenicol stock.

Plasmid Isolation and DNA quantification. Plasmids were extracted from 2 mL of overnight culture of pSPPH21-*lacI* containing DH5α cells using the PureLink Quick Plasmid Miniprep Kit (Invitrogen) and minipreps were performed following the manufacturer's instructions. A NanoDrop 2000 spectrophotometer was used to determine the DNA concentration and DNA purity of extracted plasmid DNA samples (Thermo Scientific).

Whole-Plasmid (Plasmidsaurus) Sequencing. Isolated and quantified DNA constructs, pSPPH21-*lacI*, pSPPH21:nahR/PsaI, were prepared for Plasmidsaurus sequencing according to Plasmidsaurus protocols (2.5-25kb, 30 ng/µL, > 10 µL). Alignments with the aforementioned putative sequences were performed using both SnapGene version 6.2 and EMBOSS matcher version 2.0u4 (EMBL-EBI).

Restriction enzyme digest. For construct verification of pSPPH21:nahR/PsaI plasmid vector, restriction enzymes, NcoI (NEB) and HindIII (NEB) were identified using SnapGene and BioRender. 200 ng of pSPPH21 c#1, pSPPH21 c#6.1, pSPPH21 c#13 were digested with NruI, or double digested with Ncol (NEB) and HindIII (NEB) simultaneously, according to protocols provided by NEBcloner on the NEB website. NEBuffer r3.1 was used as a digest buffer. The reaction was incubated in a 37°C water bath for 1 hour. 10 µL of each RE digestion group ran on a 1% agarose gel for gel visualization.

Agarose DNA Gel Electrophoresis. Gel electrophoresis experiments were performed using 1% agarose gels prepared in 0.5X TAE. SYBR® Safe DNA stain (Thermo Scientific) was added (1:10,000) before casting the gel. 10µL of samples each mixed with 2µL 6X Orange

gel loading dye (NEB) were loaded into the wells and 10 µL of SM1173 O'GeneRuler DNA Ladder Mix (Thermo Scientific) was loaded onto the side as a reference. Gels were run at 140 V for 90 minutes and visualized using the Bio-Rad ChemiDoc MP imaging system.

Quantification of the inducible range for pSPPH21-*lacI*. *E. Coli* DH5a containing the pSPPH21-*lacI* were obtained from Gawol *et al*. The cells were incubated for 16 hours at 37°C on a LB + chloramphenicol resistance (CmR) plate. A colony was then selected using a pipette tip and incubated into LB broth with CmR to create an overnight culture. LB was used with CmR to ensure that the cells cultured contained our vector of interest. Following incubation, the cells were pelleted and washed with 1x PBS (Sigma) to remove any background fluorescence from yeast extract. To characterize the inducible range pSPPH21-*lacI*, the cells were transferred to a 96 well plate and induced with varying concentrations of IPTG ranging from 10^{-4} µM to 10^{3}µM (11-12). Absence of cells and IPTG induction were used as negative controls. The plates were incubated for 10 hours in the BioTek Synergy H1 Multimode Microplate Reader. GFP, RFP, and OD600 readings were obtained every 3.5 minutes. For GFP signals, the excitation and emission wavelengths were recorded at 485 nm and 528 nm, respectively. For RFP signals, the excitation and emission wavelengths were recorded at 580 nm and 620 nm, respectively. Relative fluorescence units (RFU) for GFP and RFP were all normalized to OD600 to account for varying populations of cells across the wells (10-11). Data was plotted with Prism®. Standard error of the mean (SEM) was calculated by dividing the standard deviation by the sample size's square root. Statistical significance of fold changes were calculated using ordinary one-way ANOVA on Prism®.

Quantification of the inducible range for pSPPH21::nahR/PsaI. The protocol for induction of pSPPH21-*lacI* was used to analyze the pSPPH21::nahR/PsaI expression with salicylate as the inducer. The following salicylate concentrations were prepared and diluted in LB broth, to induce pSPPH21::nahR/PsaI, in triplicates, across the span of 10 hours: 0 (uninduced), 0.1, 10, 100, and 1000 µM of salicylate. Data was plotted with Prism®. Standard error of the mean (SEM) was calculated by dividing the standard deviation by the sample size's square root.

RESULTS

Whole-plasmid sequencing confirmed sequence of pSPPH21-*lac*-GFP and pSPPH21-*lac*-RFP. We set out to sequence verify the duo-directional *lac* operon containing pSPPH21 constructs, pSPPH21-*lac*-GFP and pSPPH21-*lac*-RFP by performing whole-plasmid sequencing. Plasmid maps for pSPPH21-lac GFP and pSPPH21-lac RFP were obtained and visualised using SnapGene software (Figure 1A, B). Based on the findings of Gawol *et al.*, the size of pSPPH21 plasmid without the *lac*-insert was found to be 3520 bp and the *lac*-insert was found to be 1600 bp (6). The sequencing result reveals the dominant peak in 5126 bp as shown in Figure 1 C. Similarly, we see an additional 1606 bp which should be the lac insert. The plasmid map also confirms the identity of pSPPH21-*lac*-GFP and pSPPH21-*lac*-RFP as all features are accounted for (Figure 1A, B).

Restriction enzyme (RE) digest failed to confirm the overall integrity of pSPPH21::*nahR/PsaI* construct. Previously, Frese *et al.* reported unknown deletions in the promoter and more upstream gene elements based on their Sanger sequencing results and restriction digestion experiments (12). To validate the overall identity of both pSPPH21::*nahR/PsaI*-GFP and pSPPH21::*nahR/PsaI*-RFP constructs by performing whole-plasmid sequencing and a restriction enzyme (RE) digest.

From plasmidsaurus sequencing results, the dominant peaks for pSPPH21::nahR/PsaI c#1, c#13, and c#6.1 appeared at 2645 bp, 3098 bp, and 3522 bp respectively (Figure 2B-D). All three colonies have a single dominant peak generated, however for c#1, the number of reads is significantly low (Figure 2B-D).

Using molecular biology softwares, SnapGene and BioRender, cut sites within the RFP, GFP, and in the plasmid backbone were identified and the restriction enzymes NcoI and HindIII were selected and a virtual double RE digestion was performed (Figure 2A). The digestion with NruI showed no change compared to the uncut plasmid, which is the expected result as insertion of the promoter would destroy the site (Figure 2E). The selected enzymes

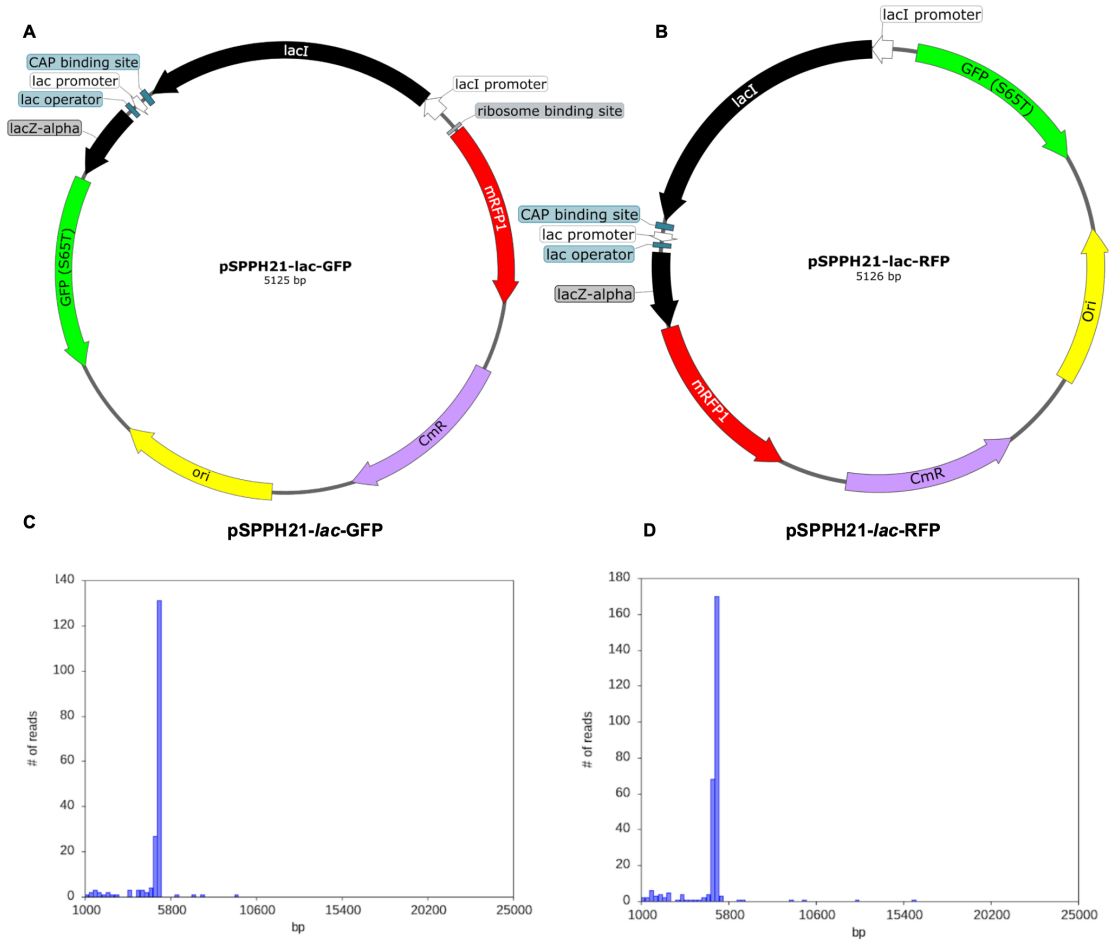

FIG. 1 Whole-plasmid Sequencing (Plasmidsaurus) results for pSPPH21-lac-GFP and pSPPH21-lac-RFP Plasmids. (A) Plasmid map for pSPPH21-lac GFP **(B)** Plasmid map for pSPPH21-lac RFP **(C)** Histogram revealing number of reads (y-axis) in respect to varying basepair (x-axis) in the lengths of all the species in pSPPH21-lac GFP plasmid. (D) Histogram revealing number of reads (y-axis) in respect to varying basepair (x-axis) in the lengths of all the species in pSPPH21-lac RFP plasmid

NcoI and HindIII cut once within the GFP gene and twice in the RFP gene respectively giving expected band sizes to be around 1045, 1520 and 3271 base pairs (bp) long (Figure 2A). While we see some faint shifts in band sizes in our double RE digest, the expected band sizes were not visible on the gel and all three pSPPH21::nahR/PsaI vectors originating from three separate glycerol stocks had similar band sizes compared to the undigested control, which were also consistent with the number of base pairs obtained from the whole-plasmid sequencing results (Figure 2B-E). This suggests that the restriction enzymes NcoI and HindIII used were unable to digest the pSPPH21::nahR/PsaI plasmid.

pSPPH21-*lacI* expression is sensitive to varying IPTG concentrations. Next, we investigated the dose-response relationship of IPTG and pSPPH21-*lacI* to provide insight into the mechanisms underlying SIGEX's response. Varying concentrations of IPTG were used to induce expression of RFP/GFP and the absence of IPTG and host strain of pSPPH21-*lacI* served as a negative control (Figure 3). Fluorescence signals were measured using the BioTek Synergy H1 Microplate Reader and normalized to OD600. The maximum RFU/OD600 readings following 10 hours of induction were log-transformed and plotted (Figure 3A, B). We observed that GFP and RFP expression appears to be positively correlated with increasing IPTG concentrations, eventually reaching around 300,000 RFU for the GFP direction and 6,000 RFU for the RFP direction. The max RFU/OD600 for RFP shows a drop off between 10^2 µM and 10^4 µM and appears to reach intensities much lower in comparison to GFP signals. Similarly, a time course assay across the span of 10 hours shows that the inducible range of pSPPH21-*lacI* lies between 1 µM to 10 µM (Figure 3C, D). The signal increases over time, however there is a decrease in the first 2 hours before increasing for the duration

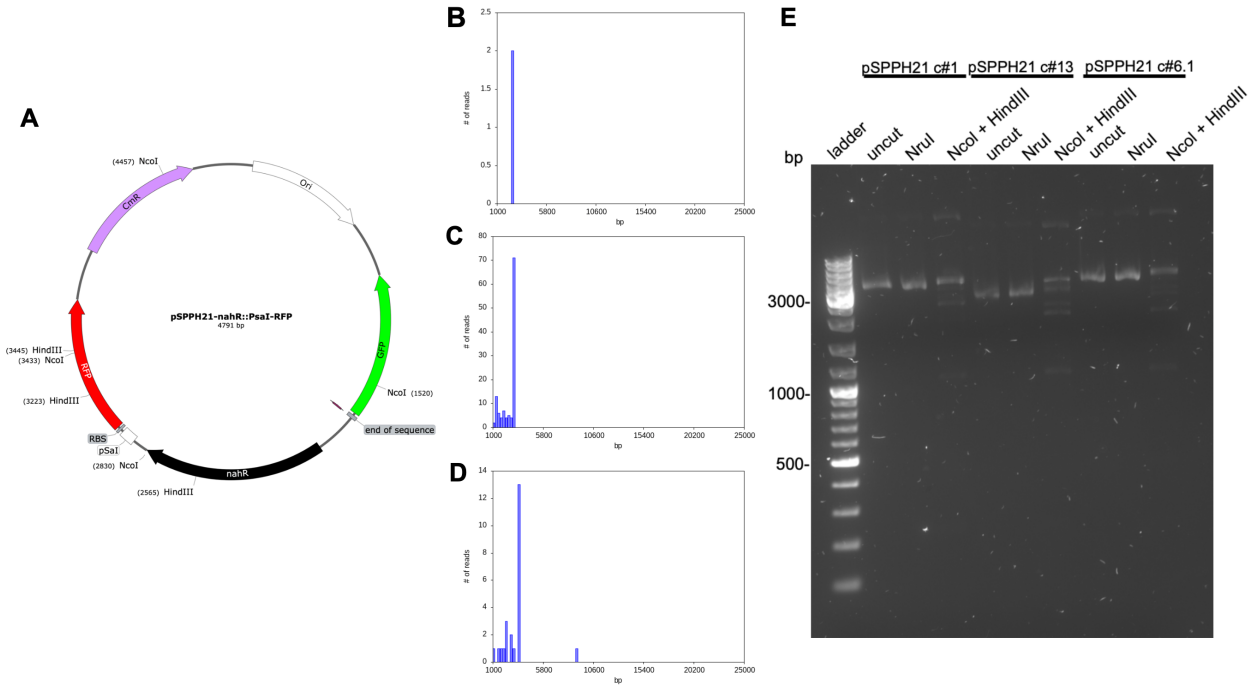

FIG. 2 Restriction digest failed to confirm the construct of pSPPH21::nahR/PsaI. Restriction digest was designed using SnapGene and BioRender. **(A)** Restriction cut sites of NruI, NcoI+HindIII on pSPPH21::nahR/PsaI. **(B)** Plasmidsaurus sequencing result for pSPPH21::nahR/PsaI c#1. **(C)** Plasmidsaurus sequencing result for pSPPH21::nahR/PsaI c#13. **(D)** Plasmidsaurus sequencing result for pSPPH21::nahR/PsaI c#6.1. **(E)** 1% gel electrophoresis of restriction digests. Lane 1: DNA ladder. Lane 2,5,8: undigested pSPPH21::nahR/PsaI. Lane 3,6,9: single digestion by NruI. Lane 4,7,10: double digestion by NcoI + HindIII. Lane 2,3,4: pSPPH21::nahR/PsaI c#1. Lane 5,6,7: pSPPH21::nahR/PsaI c#13. Lane 8,9,10: pSPPH21::nahR/PsaI c#6.1.

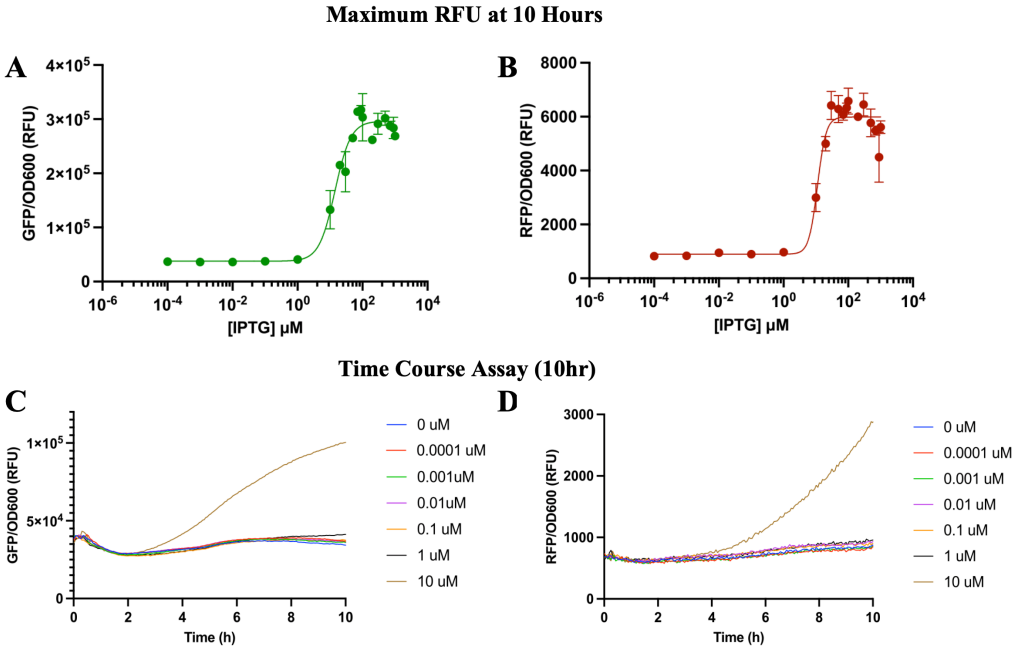

FIG. 3 pSPPH21-lacI GFP/RFP expression exhibits sensitivity to varying concentrations of IPTG. *E. Coli DH5a* cells carrying pSPPH21-lacI were induced with varying IPTG treatments ranging from 10^{-4} µM to 10^{3} µM. *E. Coli DH5a* cells were incubated in LB with the corresponding IPTG concentration and readings were obtained over the span of 10 hours. GFP/RFP signal levels were normalized to OD600 and error bars indicate standard error of mean between three replicates. pSPPH21-lacI GFP **(A)** and RFP **(B)** expression follows a dose-response for [IPTG]. The minimum concentration of IPTG that results in pSPPH21-lacI expression of GFP **(C)** and RFP **(D)** is between 1 µM to 10 µM.

of the experiment for every concentration in every plate reader including the no IPTG control. Our results show that using a microplate reader to read GFP, RFP, and OD600 signals is an efficient and high-throughput method of collecting data and that pSPPH21-*lacI* plasmid is sensitive to IPTG induction.

GFP:RFP ratio may allow for direct comparison between expression in both directions of the SIGEX construct. To further characterize GFP and RFP expression and enable the comparison between them, we examined levels of GFP and RFP fluorescence following IPTG induction of pSPPH21-*lacI* containing *E. coli* DH5a host cells. The data was obtained from the microplate experiments using a BioTek Synergy H1 Microplate Reader and then normalized to OD600. The GFP/RFP ratio from the 100 μM sample was calculated and plotted over time. Our results show that the GFP:RFP fluorescence signal does not fall below 46:1 and reaches a maximum of 109:1 at 3.25 hours of induction (Figure 4). GFP:RFP begins to increase roughly 1 hour after induction and decreases following 4 hours of induction with IPTG. We see that aside from an increase in the ratio in the middle of the experiment, the GFP:RFP ratio is at around 50 at the start and end of the experiment.

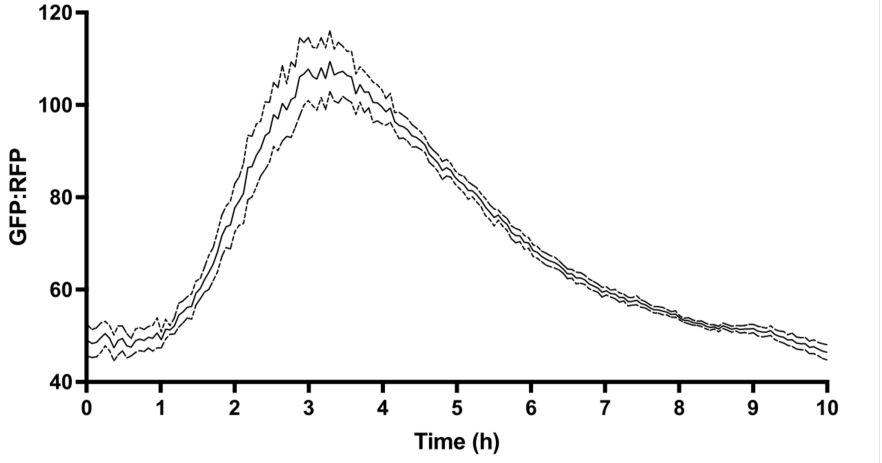

FIG. 4 GFP signal is consistently higher than RFP signal following IPTG induction of pSPPH21-lacI. *E.Coli DH5a* containing pSPPH21-lacI were induced with 100 μM of IPTG and incubated for 10 hours to measure GFP:RFP fluorescence signal in units of RFU normalized to OD600. Dotted lines indicated the SEM across six biological replicates.

Salicylate failed to induce expression of pSPPH21::nahR/PsaI plasmid. Characterization of the pSPPH21::nahR/PsaI plasmid was done with various concentrations of salicylate induction in a microplate for 10 hours overnight. DH5α cells containing pSPPH21-*lacI* were used as negative control while colonies 1, 6.1, and 13 containing pSPPH21::nahR/PsaI were obtained from Frese *et al*. Again, plate reader data was collected with a BioTek Synergy H1 Microplate Reader and fluorescent signals were normalized to OD600. We observe that unlike the *lac* operon, induction of the insert does not result in increasing signals proportional to the concentration of the inducer. Concentrations of up to 1 mM salicylate do not have an effect on normalized GFP or RFP fluorescence (Figure 5A, B) as the graph does not increase after 2 hours, and this is also the case for other colonies including the pSPPH21-*lacI* control. While analyzing the fluorescence of each sample at 10 hours, we see that there is no statistical significance between groups (Figure 5C, D). This suggests that salicylate induction of concentrations up to 1mM did not induce the expression of GFP or RFP in pSPPH21::nahR/PsaI containing *E. coli* cells. Since whole-plasmid sequencing results shows that the promoter is absent from the sequence of the pSPPH21::nahR/PsaI vector, this would prevent RNA polymerases from binding to the promoter region to facilitate transcription and expression of downstream fluorescent reporter elements (12).

DISCUSSION

We observe relatively large secondary peaks for both pSPPH21-lac GFP andpSPPH21-lac RFP plasmids in the number of reads at around 4800 bp, which indicates potential contamination or varying number of indels (Figure 1C, D). Unlike Sanger sequencing, which uses a primer to target specific sequences, Plasmidsaurus adopts a nanopore sequencing approach that sequences both the whole plasmid as well as everything in the sample. Varying numbers of indels may cause the read lengths to straddle a bin boundary and result in the

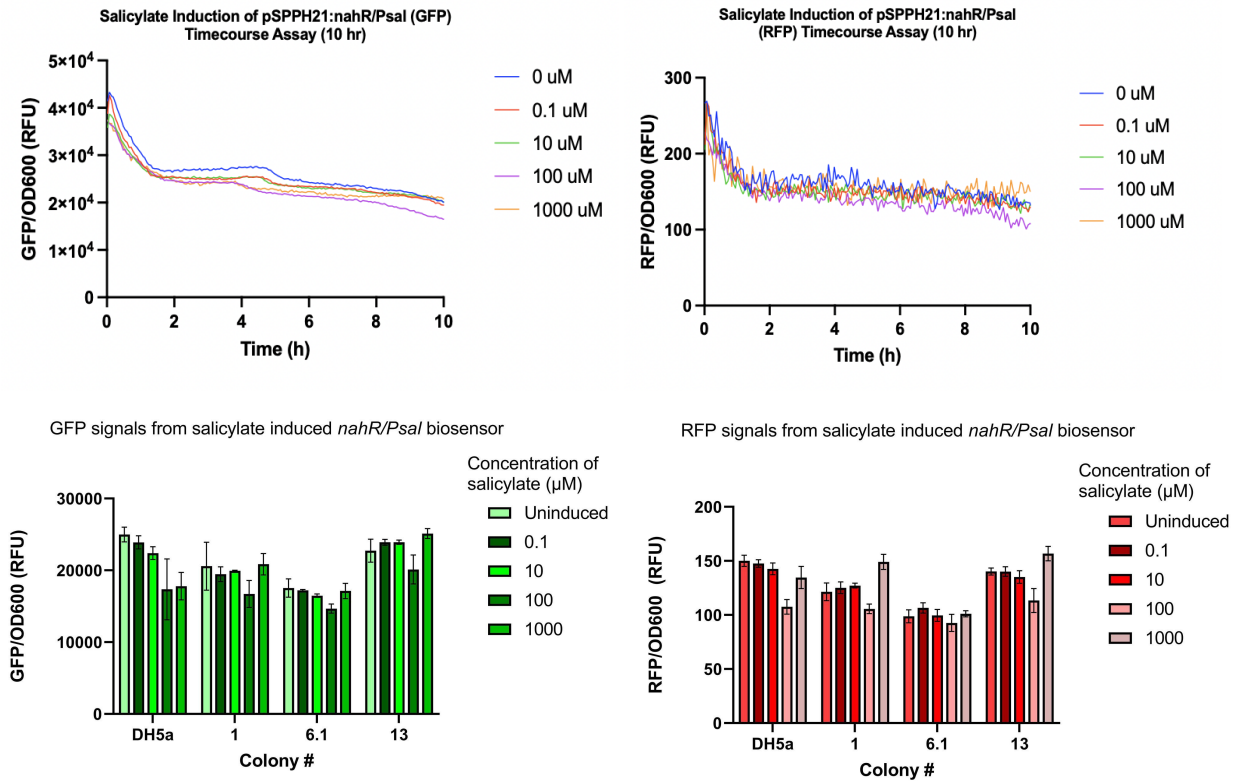

FIG. 5 pSPPH21::nahR/PsaI plasmid shows no responsiveness to salicylate inductions of concentrations up to 1 mM. *E. coli* DH5α cells transformed with pSPPH21::nahR/PsaI biosensor were induced with varying concentrations of salicylate. Cell culture was cultured overnight, washed, then incubated in LB media and salicylate for 10 hours at 37°C while shaking in a plate reader. Readings were taken every 3 minutes and normalized to OD600 and each concentration of salicylate was run in triplicates. Error bars indicate SEM. Representative **(A)** GFP and **(B)** RFP time course from colony #1 shows that pSPPH21::nahR/PsaI is not induced by salicylate. Endpoint readings at 10 hr show that there is no significant difference between concentrations for **(C)** GFP and **(D)** RFP when an ordinary one-way ANOVA is performed ($p = 0.26$ for RFP and $p = 0.37$ for GFP).

appearance of a secondary peak adjacent to the dominant peak. Therefore, these two peaks likely come from a single plasmid. Thus, we are able to ascertain the DNA sequence composition of both pSPPH21-*lac*-GFP and pSPPH21-*lac*-RFP. For future studies one can further purify the sample by reselecting more colonies from the source stock or re-ligating the *lac* inserts into the pSPPH21 plasmid to try to avoid contamination.

All three colonies of pSPPH21::nahR/PsaI give a sequencing result with one dominant peak at around 3000 bp. For colony 1, the read count is too low to distinguish any peaks or to generate any consensus. Likewise, the read count for c#13 and c#6.1 are relatively low yet distinct enough for a consensus. This is likely due to low DNA concentration during sample preparation. Based on Gawol *et al.'s* findings, the size of pSPPH21 plasmid alone was found to be 3520 bp but all three colonies of pSPPH21::nahR/PsaI are smaller. This means that there are either unknown deletions in the plasmid or the salicylate insert was not introduced properly. The gel electrophoresis result for pSPPH21::nahR/PsaI did not successfully reveal the overall integrity of pSPPH21::nahR/PsaI construct either. For the double RE digest with NcoI and HindIII, some faint shifts in bands can be observed in lane 4, 7, 10 which could be due to vectors being linearized (Figure 2E). However, expected band sizes were not visible and the brightest bands were shown in the same sizes as undigested control. This implies restriction enzymes NcoI and HindIII are unable to digest the pSPPH21::nahR/PsaI construct. As for improvements, it is imperative to include a positive control such that a restriction enzyme would have also cut in the backbone somewhere essential such as the origin to further validate restriction enzymes NcoI and HindIII are unable to digest the pSPPH21::nahR/PsaI construct. This allows differentiation between a failed digestion reaction or being uncut due

to a missing cut site. Taken together, the sequencing result and gel electrophoresis result suggest that there are deletions in all three colonies of the pSPPH21::nahR/PsaI construct.

Previous experiments by Gawol *et al*. were able to demonstrate that GFP and RFP are expressed in pSSPH21-*lacI* in response to IPTG, but the inducible range of the expression was not characterized (6). Understanding the sensitivity of the pSSPH21-*lacI* to IPTG can help determine the inducible range for future SIGEX based promoter traps. A time course assay for 10 hours was conducted to visualize the GFP and RFP expression levels of pSPPH21-*lacI* across varying concentrations of IPTG (Figure 4C, D). The time course assay for IPTG induction of pSPPH21-*lacI* showed that the minimum time to induce GFP/RFP expression with IPTG is 2 hours, which is much faster than established protocols from literature (4). The inducible range for pSPPH21-*lacI* appears to fall between 1 μM to 10 μM and any IPTG concentration below 1 μM results in control levels of expression. Therefore, we can conclude that at least 10 μM of IPTG needs to be administered to induce pSPPH21-*lacI* expression and shows that the lac promoter inserted into SIGEX is sensitive. Our data also suggests that there is little to no basal level of expression of the fluorescent protein reporter genes. The GFP and RFP signals detected for both our 0 mM IPTG group, and no cells control group are similar. However, the normalized RFU/OD600 values decrease over time as the cells grow and the RFUs remain constant for our uninduced group, while the RFU/OD600 values remain relatively consistent for the no cell control group. The inducible parameters that we identified for pSPPH21-*lacI* expression could be beneficial for future SIGEX promoter trap constructs. A small amount (10μM) of inducer has been shown to promote plasmid expression, indicating that the duo-directional SIGEX may be highly efficient in identifying novel promoters

To utilize pSPPH21 as a duo-directional promoter-discovery tool in environmental metagenomic applications, it is necessary to characterize the fluorescence intensity of GFP and RFP expressed by GFPmut3 and mRFP1, respectively. Following IPTG induction of pSPPH21-*lacI*, we observed that the GFP:RFP signal can range from approximately 50:1 to 100:1 across the 10 hour induction period (Figure 3). This may be attributed to the difference in molecular structure and therefore light absorption and emission between mRFP1 and GFPmut3. Since the GFP:RFP signal following IPTG induction does not remain consistent throughout a 10 hour time span, it may be difficult to draw direct comparisons. GFP:RFP starts to increase after 1 hour, reaching a maximum signal ratio, and subsequently decreases at4 hours. This may be due to the intrinsic differences in GFP and RFP intensities as GFP signal in response to the same IPTG concentration results in larger RFUs than RFP signal. However, our data shows that after around 7 hours, the ratio curve begins to flatten out. Indeed, Figure 2A suggests that the maximum RFUs in duo-directional SIGEX is around 300,000 RFU for GFP and 6,000 RFU for RFP in response to IPTG induced lac promoter, as IPTG concentrations greater than 1 mM have little effect on signal increase. From Figure 4, there may be merit in performing SIGEX induction experiments for at least 10 hours to obtain RFUs that can be directly compared. A potential concern is that, from Figure 3, we see that GFP (Figure 3C) begins to plateau towards the end of the experiment while RFP (Figure 3D) appears to still be increasing exponentially. This may be due to changes in expression of sigma factor Sigma 70 (Sig70) promoters during stationary phase as a result of cells experiencing stress due to nutrient deprivation or perhaps a limitation of the plate reader as the signal strength approaches saturation levels. As such, the ability to compare expression levels and therefore promoter induction may be hindered by this observation. Further work with other inducible promoters could be done to explore the GFP and RFP ratio to be able to draw qualitative comparisons and to validate the potential 50:1 ratio we see in the lac promoter.

Our IPTG induction results show similar sensitivity to that of other papers (14). Also notable is the slight decrease in signal for both GFP and RFP as the concentration of IPTG increases past 0.1 mM, which is likely due to the toxic nature of IPTG (13). However, we see that the decrease is slight, up to the normally recommended concentration of 1 mM, suggesting that using large amounts of inducer to achieve a signal is acceptable.

Performing the experiment in a microplate provides proof-of-concept for further metagenomic work employing duo-directional SIGEX. As our microplate protocol collects OD600 data in addition to GFP and RFP, this allows for the normalization of data and

therefore a high throughput method of metagenomic promotor screening. A metagenomic sample could be loaded into each well, allowing for 96 inducers to be tested overnight at a time. We also observe that the first 2 hours of induction show a decrease in fluorescent signal for both the induction of *lac* promoter and *nahR/PsaI*. This may be attributed to the lag phase, where cells are adjusting to different environments as well as the degradation of any existing GFP by cellular proteases. It may also be due to a lag phase in the expression of reporter genes, leading to an increased OD600 reading and therefore a lower normalized fluorescent signal (9, 16). However, we also observe this dip for every concentration tested, including the no IPTG control group. Perhaps the growth rate of the bacteria is initially greater than the expression rate of GFP/RFP, decreasing the OD600/RFU value until the growth rate of the bacteria decreases to be less than the expression rate of GFP, leading to an increase in OD600/RFU. Indeed, β-galactosidase activity increases almost immediately after IPTG induction (17). Thus, our time course assay suggests that induction of SIGEX plasmids should be done for at least 4 hours to bypass an observed dip in normalized fluorescent signal. Additionally, as the fluorescent signal is maintained and continues to increase, induction of SIGEX plasmids for 10 hours or more may be ideal. Further work can be done to validate that SIGEX is sensitive for many types of inducers to ensure its broad application in metagenomic screening.

We selected our salicylate concentrations based on work done by the 2013 Peking iGEM team (https://2013.igem.org/Team:Peking/Project/BioSensors/NahR). They were able to observe the fluorescent signal of sfGFP with concentrations as low as 1 μM. Indeed, other iGEM teams reported similar findings, where the nahR/PsaI biosensor fluoresces with concentrations less than 1μM of salicylate (http://parts.igemPart:BBa J61051.org/Part:BBa_J61051). Our results from the pSPPH21::nahR/PsaI salicylate induction assay are similar to those of Frese *et al*. and we see no fluorescence up to 1 mM of salicylate induction (10). In addition, our sequencing results and double digest shows issues with plasmid integrity. Sequencing results show large segments of deletions in all psPPH21::nahR/PsaI plasmids. Failure to digest pSPPH21::nahR/PsaI plasmid with NcoI and HindIII can be attributed to the unsuccessful insertion of the *nahR/PsaI* insert, and therefore the lack of sites for the restriction enzymes to target. However, previous studies by Frese et al. showed that induction of pSPPH21::nahR/PsaI can only be seen after inducing with a minimum of 10 mM salicylate (10). We decided that there was no merit in re-inducing the plasmid with higher concentrations of salicylate as our sequencing and digest results indicate that there are multiple deletions in the plasmid. A more efficient follow up experiment would be to attempt another insertion of the nahR/Psal biosensor, as previous iGEM teams have shown that they were able to successfully clone it into plasmid.

Conclusions This work contributes to our understanding of the duo-directional SIGEX plasmid by further characterizing its functionality. We verified the function of pSPPH21-*lacI* and showed that it is sensitive to various concentrations of IPTG. By performing a time course assay of GFP:RFP, we were able to observe that 100 μM of inducer increased the GFP:RFP signal proportionally between 2 to 4 hours and therefore suggest avoiding analyzing data in the first 2-4 hours of induction. Additionally, there is the opportunity to examine the GFP:RFP ratio further in order to be able to compare the inducibility of the promoter regardless of GFP and RFP directionality when duo-directional SIGEX is applied to metagenomics. The pSPPH21::nahR/PsaI plasmid needs to be reinserted as our sequencing results show segments of deletions and concentrations of up to 1 mM salicylate cannot induce GFP or RFP expression.

Future Directions The purpose of duo-directional SIGEX is to be able to detect an inducible promoter ligated in either direction. However, the fluorescence of GFP is greater than RFP, making comparison difficult. Future experiments analyzing different inducible promoters to determine if the GFP to RFP ratio is consistent for certain concentrations of inducer can allow us to be able to quantify the amount of expression regardless of directionality.

The next step is to use a metagenomic sample and test the functionality of our SIGEX plasmid in identifying novel promoters in the presence of inducible substrates. Furthermore, one could test whether different promoters across different taxonomic species, for instance

Mycobacterium tuberculosis, behave similarly to the lac promoter in *E. coli.* And ultimately, redesigning the salicylate SIGEX vector (pSPPH21::*nahR/pSaI*) by adding a Kanamycin resistance gene for double antibiotic selection.

ACKNOWLEDGEMENTS

We acknowledge that the land we carry this project is situated on the traditional, ancestral, and unceded territory of the xwməθkwəy̓əm (Musqueam) People. This project was funded by the Department of Microbiology and Immunology at the University of British Columbia. We thank Dr. David Oliver, Jade Muileboom, and Gara Dexter for their mentorship, technical advice, and constant support throughout this project. Their feedback and guidance were invaluable throughout the experimental design and troubleshooting process. We further thank Frese *et al.* and Gawol *et al.* for providing the pSPPH21-*lacI* and pSPPH21::nahR/PsaI plasmids.

CONTRIBUTIONS

All participants participated equally in the project and in writing the manuscript. All four members contributed equally to the experiments. R.J.H conceptualized the study and all four members conducted background research and optimized the study. All four members contributed equally to data analysis. J.Z., G.W., and R.J.H. contributed towards designing, planning and troubleshooting of plate reader experiments. T.Z. and G.W. performed gel electrophoresis experiments. R.J.H. wrote the abstract, introduction, results, future directions, and discussion, and analyzed sequencing results and performed all plasmid isolation experiments from *E. coli*. G.W. designed the figures and wrote the methods, results, and discussion for all pSPPH21-lacI induction related experiments. J. Z. generated figures and wrote the abstract, conclusion, methods, results, and discussion for all pSPPH21:nahR/PsaI induction related experiments. T.H.Z generated figures and wrote the abstract, methods, results and discussion for sequencing and gel electrophoresis, future directions, and references. R.J.H, J.Z., G.W. and T.H.Z edited the manuscript.

REFERENCES

1. **Uchiyama T, Watanabe K. 2008.** Substrate-induced gene expression (SIGEX) screening of metagenome libraries. *Nat Protoc* **3**:1202–1212
2. **Uchiyama T, Abe T, Ikemura T, & Watanabe K. 2005.** Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes. *Nat Biotechnol* **23**:88–93
3. **Meier MJ, Paterson ES, & Lambert IB. 2016.** Use of substrate-induced gene expression in metagenomic analysis of an aromatic hydrocarbon-contaminated soil. *Appl Environ Microbiol* **82**:897–909
4. **Zaslaver A, Bren A, Ronen M, et al. 2006.** A comprehensive library of fluorescent transcriptional reporters for Escherichia coli. *Nat Methods* **3**:623–628
5. **Abrishamkar P, Dana HM, Oveisi S, & Tabassi P. 2022.** Construction of SIGEX duo-directional reporter plasmid pSPPH21. *UJEMI+* **8**:1-11
6. **Gawol, D., Floyd, R., Kohara, K., & Lee, Y. 2022**. Duo-directionality of the substrate-induced expression screening vector pSPPH21 confirmed with a lac operon screen. *Undergraduate Journal of Experimental Microbiology and Immunology (UJEMI)* **8**:1-12
7. **Agnew S, Bangayan A, Fan J, & Lin A. 2022**. Testing the functionality of SIGEX duo-directional reporter plasmid pSPPH21 using an inducible promoter. *Undergraduate Journal of Experimental Microbiology and Immunology (UJEMI)* **27**:1-15
8. **Yu TC, Liu WL, Brinck MS et al. 2021.** Multiplexed characterization of rationally designed promoter architectures deconstructs combinatorial logic for IPTG-inducible systems. *Nat Commun* **12**:325
9. **Stevenson K, McVey A, Clark I et al. 2016.** General calibration of microbial growth in microplate readers. *Sci* Rep **6**, 38828
10. **Astral Scientific. 2022.** IPTG: Triggers the transcription of the Lac Operon. *Astral Scientific*. Retrieved October 12, 2022, from https://astralscientific.com.au/blogs/news/iptg-triggers-the-transcription-of-the-lac-operon
11. **BiologicsCorp. 2018**. IPTG induction protocol. *BiologicsCorp.* Retrieved October 12, 2022, from https://www.biologicscorp.com/blog/iptg-induction-protocol/#.Y0b5vXbMKUk
12. **Frese K., Martens, K., Nichvolodoff, T. 2022.** Multiple deletions arise from attempted insertion of the salicylate nahR/Psal biosensor into a duo-directional SIGEX vector. *Undergraduate Journal of Experimental Microbiology and Immunology (UJEMI)* **27**:1-14

13. **Rizkia, P. R., Silaban, S., Hasan, K., Kamara, D. S., Subroto, T., Soemitro, S., Maksum, I. P. 2015.** Effect of Isopropyl-β-D-thiogalactopyranoside Concentration on Prethrombin-2 Recombinant Gene Expression in Escherichia Coli ER2566, Procedia Chemistry. **17**, 118-124
14. **Mühlmann, M., Forsten, E., Noack, S., & Büchs, J. 2017.** Optimizing recombinant protein expression via automated induction profiling in microtiter plates at different temperatures. *Microbial cell factories*, **16**(1), 220.
15. **De Lorenzo V. 2005.** Problems with metagenomic screening. *Nature Biotechnology* **23**:1045–1045.
16. **Wilson, C. J., Zhan, H., Swint-Kruse, L. & Matthews, K. S. 2007.** The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. *Cell. Mol. Life Sci.* **64**, 3–16.
17. **Marbach, A. & Bettenbrock, K. 2012.** lac operon induction in Escherichia coli: Systematic comparison of IPTG and TMG induction and influence of the transacetylase LacA. *Journal of Biotechnology.* **157**(1) 82-88.
18. **Helmann, J. D. & Chamberlin, M. J. 1998.** Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.* **57**, 839–872
19. **Müller, J., Oehler, S. & Müller-Hill, B. 1996.** Repression of lac promoter as a function of distance, phase and quality of an auxiliary lac operator. *J. Mol. Biol.* **257**, 21–29
20. **Oehler, S., Amouyal, M., Kolkhof, P., von Wilcken-Bergmann, B. & Müller-Hill, B. 1994.** Quality and position of the three lac operators of E. coli define efficiency of repression. *EMBO J.* **13**, 3348–3355
21. **Becker, N. A., Peters, J. P. & Maher, L. J. 2013.** Mechanism of promoter repression by Lac repressor-DNA loops. *Nucleic Acids Res.* **41**, 3962–3962
22. **Camsund, D., Heidorn, T. & Lindblad, P. 2014.** Design and analysis of LacI-repressed promoters and DNA-looping in a cyanobacterium. *J. Biol. Eng*. **8**, 4
23. **Brewster, R. C., Jones, D. L. & Phillips, R. 2012.** Tuning promoter strength through RNA polymerase binding site design in Escherichia coli. *PLoS Comput. Biol*. **8**, e1002811