Discovery and Characterization of Native *Deinococcus radiodurans* Promoters for Tunable Gene Expression

Angela Chen,[a] Mark W. Sherman,[b] Cynthia Chu,[a] Natalia Gonzalez,[a] Tulshi Patel,[a] Lydia M. Contreras[a,b]

[a]McKetta Department of Chemical Engineering, University of Texas at Austin, Austin, Texas, USA
[b]Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, Texas, USA

ABSTRACT The potential utilization of extremophiles as a robust chassis for metabolic engineering applications has prompted interest in the use of *Deinococcus radiodurans* for bioremediation efforts, but current applications are limited by the lack of availability of genetic tools, such as promoters. In this study, we used a combined computational and experimental approach to identify and screen 30 predicted promoters for expression in *D. radiodurans* using a fluorescent reporter assay. The top eight candidates were further characterized, compared to currently available promoters, and optimized for engineering through minimization for use in *D. radiodurans*. Of these top eight, two promoter regions, *PDR_1261* and *PrpmB*, were stronger and more consistent than the most widely used promoter sequence in *D. radiodurans*, *PgroES*. Furthermore, half of the top eight promoters could be minimized by at least 20% (to obtain final sequences that are approximately 24 to 177 bp), and several of the putative promoters either showed activity in *Escherichia coli* or were *D. radiodurans* specific, broadening the use of the promoters for various applications. Overall, this work introduces a suite of novel, well-characterized promoters for protein production and metabolic engineering in *D. radiodurans*.

IMPORTANCE The tolerance of the extremophile, *Deinococcus radiodurans*, to numerous oxidative stresses makes it ideal for bioremediation applications, but many of the tools necessary for metabolic engineering are lacking in this organism compared to model bacteria. Although native and engineered promoters have been used to drive gene expression for protein production in *D. radiodurans*, very few have been well characterized. Informed by bioinformatics, this study expands the repertoire of well-characterized promoters for *D. radiodurans* via thorough characterization of eight putative promoters with various strengths. These results will help facilitate tunable gene expression, since these promoters demonstrate strong and consistent performance compared to the current standard, *PgroES*. This study also provides a methodology for high-throughput promoter identification and characterization using fluorescence in *D. radiodurans*. The promoters identified in this study will facilitate metabolic engineering of *D. radiodurans* and enable its use in biotechnological applications ranging from bioremediation to synthesis of commodity chemicals.

KEYWORDS *Deinococcus radiodurans*, promoter engineering, biotechnology, extremophile, regulatory systems

Address correspondence to Lydia M. Contreras, lcontrer@utexas.edu.

A.C. and M.W.S. contributed equally to this article.

Recent work regarding the biological engineering of extremophiles has increased interest in their use as a robust chassis for metabolic engineering. The appeal of using extremophiles in various applications is largely due to their ability to survive conditions toxic to traditional engineering strains. One extremophile considered attractive is *Deinococcus radiodurans*, a Gram-positive bacteria known for its tolerance to ionizing radiation, heavy metal exposure, desiccation, UV radiation, oxidizing agents, and electrophilic mutagens (1–3). In the context of its applications in bioremediation,

multiple efforts have been made to overexpress heterologous proteins in *D. radiodurans*, resulting in the reduction of mercury(II), toluene, chlorobenzene, 3,4-dichloro-1-butene, indole, and uranium under chronic irradiation levels similar to those experienced at radionuclide-contaminated bioremediation sites (2, 4–6). Efforts have also focused on tuning expression levels of natively occurring regulatory RNAs to construct more robust strains (7). Despite the promise for genetically engineered *D. radiodurans* in applications, its complex biology and the lack of available genetic tools, such as promoters, have hindered the ability to engineer industrial strains (5, 8–11).

Promoters used in *D. radiodurans* have traditionally been identified by their proximity to downstream genes of interest or by their anticipated response to stress and then characterized by their ability to drive expression of a reporter gene, such as β-galactosidase. Examples include the putative promoter regions of *recA*, *lexA*, *mutL*, and *recQ* (12–15). Further optimization and characterization has been performed for the constitutive promoter of *groES* by minimization, resulting in more consistent gene expression across a range of conditions compared to the full-length promoter, as well as biological insight into 5′ regulation of gene expression in *D. radiodurans* (15). The minimized 236-bp *groES* promoter is now routinely used in *D. radiodurans* for fundamental characterization of metabolic pathways, biological responses to radiation and desiccation, and regulatory pathways including sRNA interactions (2, 10, 16–20). Additional investigations attempting to characterize genetic elements in *D. radiodurans* have led to identification of three other strong constitutive promoters (*PresU*, *PtufA*, and *PtufB*) and incorporation of *Pspac*, a strong inducible promoter originating from *Bacillus subtilis* (12). The *Pspac* promoter has since been used in *D. radiodurans* for inducible gene expression and has led to progress in the identification of gene function (12, 19, 21–24).

To incorporate multiple genes and pathways into *D. radiodurans* for effective metabolic engineering applications, it is desirable to develop a suite of well-characterized promoters of various strengths (25). Although the examples described above have been utilized to drive gene expression, most remain poorly characterized relative to each other and possess a limited range of strength compared to promoters developed for expression in model species such as *Escherichia coli* and *Bacillus subtilis*. Recent advances in cloning technologies have increased the feasibility of utilizing nonmodel organisms with expanded stress tolerance for biotechnological applications. However, plasmids designed for nonmodel organisms are generally shuttled through model organisms, such as *E. coli*, for plasmid construction or amplification before being transformed into their target species. Shuttling the promoter into another system could result in unintended expression and elicit toxicity, potentially reducing cloning efficiency. Conversely, promoters functional in two organisms (i.e., *D. radiodurans* and *E. coli*) could be utilized for expression in multiple species, reducing the need to tailor every construct to a specific host organism.

In this work, we took advantage of the increasing availability of bioinformatics data, such as genomic and transcriptomic sequencing data, to enable an informed approach for the identification of novel candidate promoter sequences in *D. radiodurans* (26, 27). Specifically, we used a combined bioinformatics and experimental approach to predict and assess 30 candidate promoter sequences for their ability to promote gene expression in *D. radiodurans* as measured by a fluorescence-based assay. Of these 30 candidates, 8 functional and strong promoters were biochemically characterized and compared to others previously used in the literature. The putative promoters drove gene expression to reporter levels ranging from 2- to 15-fold greater than the background (as measured by an empty vector control). Several promoter candidates showed consistent performance regardless of growth phase. We also demonstrated minimization of four of the promoters to sizes between 24 and 177 bp and found that while some promoters were functional in both *E. coli* and *D. radiodurans* and could be conveniently used in shuttle vector systems, many were only expressed in *D. radiodurans*.
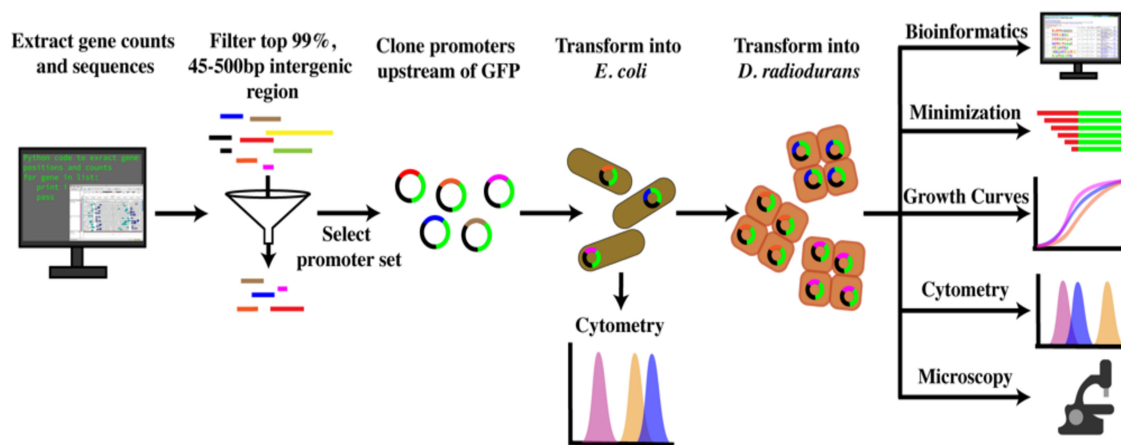
FIG 1 Bioinformatic and experimental pipeline used in this study for the identification and characterization of novel promoters. Transcriptomic and genomic data sets were used to identify intergenic regions potentially containing promoter regions in *D. radiodurans*. Regions were then filtered and cloned into a GFP reporter construct. Plasmids containing candidate promoters were transformed into *E. coli*, isolated by miniprep, and transformed into *D. radiodurans*. Regions driving reporter expression were further investigated by minimization, growth curves, cytometry, microscopy, and assessment by current bioinformatic prediction programs toward their identification.

RESULTS AND DISCUSSION

Selection of promoter candidates. We identified promoter candidates by transcriptomic analysis of gene sequences based on their transcripts per million (TPM) and then filtered by the length of the upstream intergenic region potentially containing the candidate promoter (see Data Set S1 in the supplemental material). A schematic of the screening method used to select all promoter candidates is depicted in Fig. 1. Promoter candidates were selected across a range of TPM values in an attempt to characterize promoters of various strengths. Primers used to amplify the candidate promoter regions can be found in Table 1. For brevity, putative promoter construct labels are represented by their downstream gene names in figures and in the table.

Preliminary screen of putative promoters in *D. radiodurans* and *E. coli*. To identify novel promoters in *D. radiodurans*, we assessed the exponential-phase performance of 30 strains, each carrying a single bioinformatically identified candidate promoter relative to *PgroES* (positive control) and pRadGro (empty vector negative control [EV]). We used the average median fluorescence to compare activity across candidate strains after normalization to EV, as seen in Fig. 2A. Eleven of the thirty candidate strains exhibited a significantly higher normalized average fluorescence than did EV, suggesting that the screen was able to discern true promoters from nonfunctional sequences. Interestingly, the strain carrying the promoter of *DR_1261* exhibited significantly higher fluorescence than the strain carrying *PgroES*, the most frequently used promoter for constitutive protein expression in *D. radiodurans*. Strains harboring three other promoters, *PrpmB*, *PdnaK*, and *PDR_1348*, show expression levels comparable to that of the *PgroES* control strain. In addition to these strong promoters, seven other putative promoters were identified, with strengths ranging from approximately 2- to 5-fold greater than that of EV. It is also worth noting that the measured range of fluorescent signal driven by these promoters was between 2- and 15-fold stronger than that of EV, demonstrating the synthetic utility of these newly characterized promoters. Of these candidate promoter regions, we selected the top eight candidates (25%) based on fluorescence signal for further characterization, since these were most likely to contain true promoter regions.

In addition to identifying promoter candidates in *D. radiodurans*, we evaluated the full panel of 30 candidate promoter sequences for their ability to drive green fluorescent protein (GFP) expression in *E. coli* to characterize the potential use of these promoters for applications in other organisms (Fig. 2B). Some promoters, such as *PdnaK* and *PgroEL*, demonstrated high reporter signal in both *E. coli* and *D. radiodurans*, suggesting that they could be used as potential shuttle promoters in both organisms.

TABLE 1 Primers used in this study[a]

Gene	Primer sequence (5′–3′) Forward	Reverse	Size (bp)
DR_1261	ataCCGCGGTTTCAGCATggaagttaccctaagtcc	ataCTCGAGgacctgccctgggag	116
1261-12-F	ataCTCGAGccacgcaccatc	DR_1261-R	101
1261-22-F	ataCTCGAGccatctctaacgacgcc	DR_1261-R	94
1261-32-F	ataCTCGAGcgacgccttgatgttg	DR_1261-R	84
1261-41-F	ataCTCGAGatgttgctatacgccg	DR_1261-R	74
recA	aaaCTCGAGgccggttgccgtaaa	aaaCCGCGGTTTCAGCATaacaacctcaccg	528
rpmB	aatCCGCGGTTTCAGCATgttggtactctccttgtttg	aatCTCGAGcggccacatgacgac	284
katA	aggCTCGAGcctgcgccttcttctttg	ataCCGCGGTTTCAGCATacactctccttcgcctc	211
tufB	aatCCGCGGTTTCAGCATgagtctttcctcctgg	ttaCTCGAGgcgagttcgagtctcgtt	265
dnaK	tttCCGCGGTTTCAGCATgtgttgactccttgggaaaga	attCTCGAGgcttttcttgatgttgccagc	168
groES	ttttCTCGAGgtattgtcgccctacatatatacgtt	tttCCGCGGTTTCAGCATgtggggtc	256
DR_1348	ataCCGCGGTTTCAGCATggacttatcgtgcgc	ataCTCGAGttgcttgagcgccg	110
rplL	ataCCGCGGTTTCAGCATgtgtatgtcctccagaagtg	gatCTCGAGcagcgagagcgcctaa	106
rplL-11F	aatCTCGAGgcctaaagcgcccg	rplL-R	96
rplL-20F	tttCTCGAGcccgcttctccttcc	rplL-R	86
rplL-30F	cttCTCGAGccttccctgcatttctcttt	rplL-R	77
rplL-40F	CTCGAGatttctctttccctaaaatccact	rplL-R	64
rplL-45F	aaaCTCGAGtctttccctaaaatccacttctg	rplL-R	62
rplL-50F	aaaCTCGAGccctaaaatccacttctggag	rplL-R	57
rplL-55F	aaaCTCGAGaaatccacttctggaggaca	rplL-R	52
rplL-60F	aaaCTCGAGcacttctggaggacatacacat	rplL-R	47
groEL	atttCTCGAGggctccgagtcaggttct	taaCCGCGGTTTCAGCATtgtgattgctccttgaaaagt	102
groEL-10-F	tacgccaagctcgcgaggcCTCGAGcaggttctgagcctgttcg	groEL-R	107
groEL-20-F	tacgccaagctcgcgaggcCTCGAGgcctgttcgtttcctgttt	groEL-R	97
groEL-30-F	tacgccaagctcgcgaggcCTCGAGttcctgttttttcttcctcatttc	groEL-R	87
groEL-40-F	tacgccaagctcgcgaggcCTCGAGtcttcctcatttcacttttcaag	groEL-R	77
DR_1473	aaaCTCGAGgaggaggagtctgacttgct	aaaCCGCGGTTTCAGCATggtggtgcctccttac	359
DR_2508	attatCCGCGGTTTCAGCATacgtcctccgtgagc	atGAGCTCgaggtaattctggattggcc	463
DR_2068	aatCCGCGGTTTCAGCATgctcgtgtgcctcc	cccCTCGAGccgccggctgctac	97
DR_0325	attCCGCGGTTTCAGCATgggtatccctcctgaggatgg	atCTCGAGgcttttctagcatcctggccc	212
DR_1114	cccCTCGAGgtcataaaacttgacatggat	aaaCCGCGGTTTCAGCATgtagtcctatattaaaacttgagtgc	74
AmyE	aaaCTCGAGcaccagaaggcgacgat	ataCCGCGGTTTCAGCATggggggagatgctag	217
DR_1105	aaaCTCGAGagcatttgacaaaaagatgg	aaaCCGCGGTTTCAGCATtctctgttctgctccg	118
recQ	aatCCGCGGTTTCAGCATctccccaggatagcg	aaaCTCGAGgaagcctccagatcaaagg	211
DR_1994	ataCCGCGGTTTCAGCATgccttcaatacgcg	ataCTCGAGctcagggcggc	77
rspR	attttCTCGAGtgcaggccctggggg	atttCCGCGGTTTCAGCATgttgggtgtccttggtgtgg	86
uvrA	attCCGCGGTTTCAGCATgaagctccttttggctgtcct	atCTCGAGgggtgatcgaggttgaacatcg	228
DR_2635	attCCGCGGTTTCAGCATagggtctccagaacaagctttt	attttCTCGAGtaaagccccgccttccg	141
DR_1927	ataCCGCGGTTTCAGCATctgctcagcctaaccg	ataCTCGAGgcgcttttctcccttca	233
DR_2263	cgatCCGCGGTTTCAGCATgaacgcaccctacgccg	attCTCGAGagaaaaatggctatgggcgcc	77
DR_0911	attCCGCGGTTTCAGCATtcaggctcctgtggaatggag	atttttCTCGAGtgcgggcactgccaacc	162
DR_2330	ataCTCGAGagcatttgacaaaatgatg	ataCCGCGGTTTCAGCATtctcacccatcgctgc	125
DR_1654	atttCTCGAGcgcctgctctctacgcc	atttCCGCGGTTTCAGCATtctgcaccccgaaactgac	134
DR_1689	atttCCGCGGTTTCAGCATtgcactttcccgtcagactg	aCTCGAGccttgcggaaaagttcagacttg	206
mutL	aaaCTCGAGcgtatttgccgggat	ataCCGCGGTTTCAGCATtctgttgtgagcatatcag	452
DR_1263	ataCCGCGGTTTCAGCATgttgtgtgtcgtcctca	ataCTCGAGgagcaaaatatccaaggcga	159
DR_2439	ataCCGCGGTTTCAGCATaggaatttatgttctcctcgg	ataCTCGAGgctaataggcgagtcagcc	102
LexA	aatCCGCGGTTTCAGCATgggtcaagcggcg	agaCTCGAGgagcgctacgccttcat	211
DR_0781	aaaCTCGAGcctagcctgggcgct	attCCGCGGTTTCAGCATgcgatgcccttcaaag	188
DR_2588	ataCTCGAGcgtccggggctatactcc	ataCCGCGGTTTCAGCATaaatgagaggggcagcg	143
clpB	aaaCTCGAGtatgccgtcacctggc	ataCCGCGGTTTCAGCATgtttattgaggctttaggaaagtgg	423
DR_1910	aaaCTCGAGatcatcgtcggctgagc	aaaCCGCGGTTTCAGCATcacgacgatgtgcc	270
rpml	attCCGCGGTTTCAGCATggagagccctccttcgta	attttCTCGAGcggcggctgacccc	204
DR_0907	attttCTCGAGgaaggcccgcttccca	ataCCGCGGTTTCAGCATatttggccggcgctt	135
pRad-BB	attcCCGCGGatgagcaag	attaCTCGAGgcctcgcg	7,507
pRad-BB for GA	gacgtcatatggatccgattc	ctcgaggcctcgcg	6,752
GFP colony	atggtcgacatgaccatgattacgcca	cagatgaacttcagggtcag	Variable

[a]Forward and reverse primers were used to amplify promoter inserts from *D. radiodurans* genomic DNA (NC_001263). pRadBB forward and reverse primers were used to amplify the backbone for promoter cloning. GFP colony forward and reverse primers were used to assess insertion of the promoter sequence into the backbone. Capital letters indicate cut sequences for XhoI or SacII with the codons for the first three amino acids of groES (Met-Leu-Lys coded by TACGACTTT, followed by SacII). The expected sizes of the PCR products are indicated.
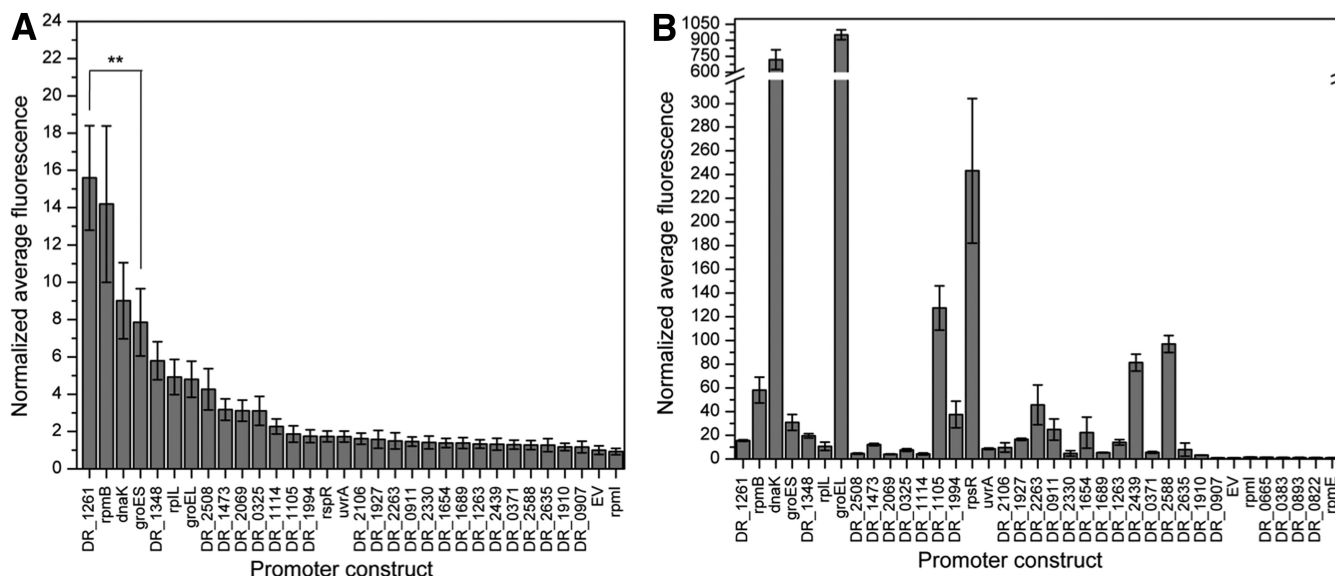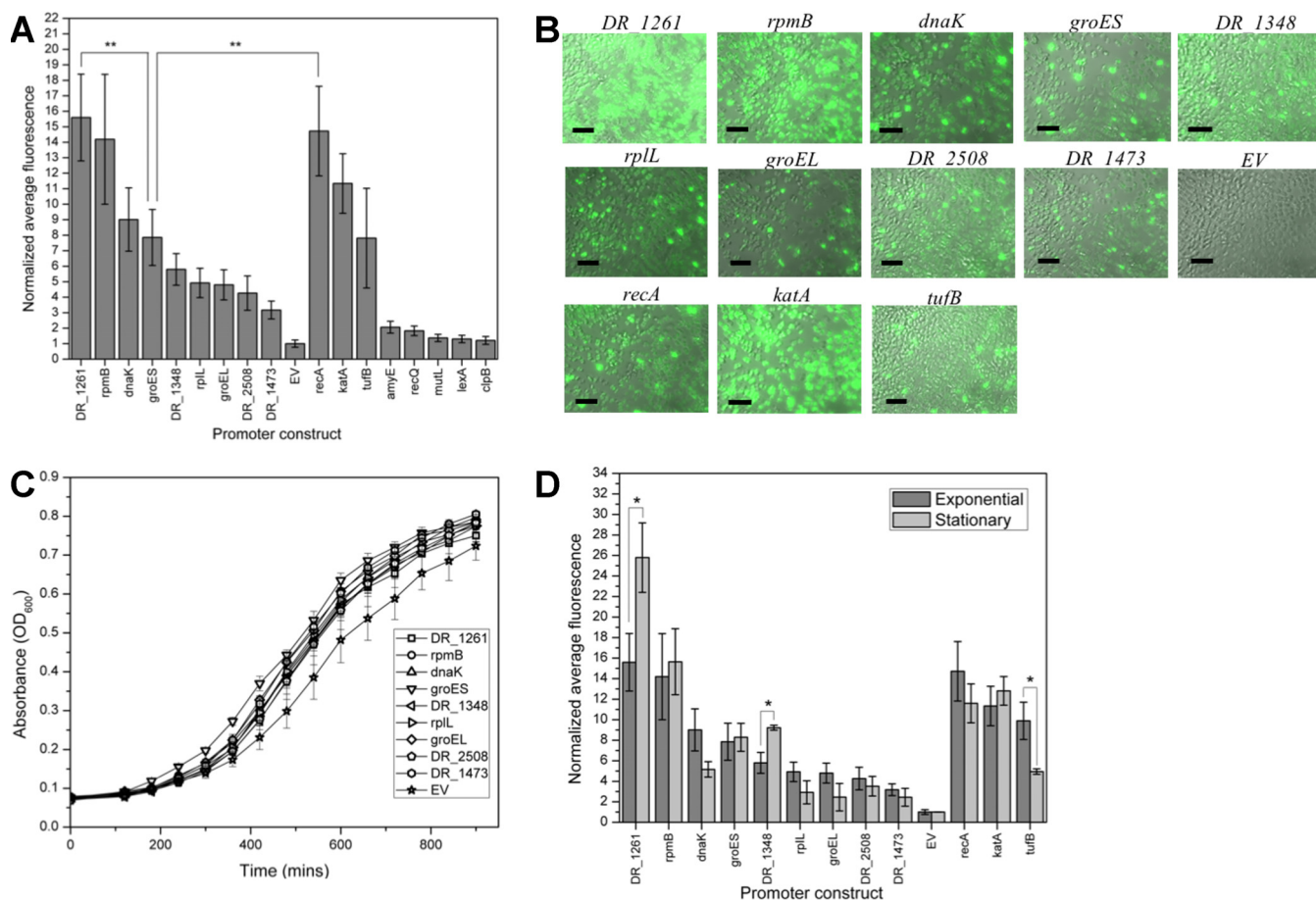
FIG 2 Normalized average median fluorescence of promoter candidates. Data were collected in biological triplicates for exponential-phase cultures grown to an OD$_{600}$ of 0.3 to 0.8 and assessed by flow cytometry. Fluorescence values, measured in arbitrary units, were normalized by EV, and error bars represent propagated standard deviations. Candidates are arranged by their strength in *D. radiodurans*. (A) Fluorescence data for candidate promoters in *D. radiodurans*. Statistical significance is denoted by asterisks (**, $P < 0.01$), as determined by Student's *t* test. (B) Fluorescence data for candidate promoters in *E. coli*.

The downstream regions of both of these promoter candidates encode known chaperone proteins, so conservation of this functional role could explain why these sequences are active in both species. Other promoter candidates, such as *PrpsR* and *PDR_1105*, show relatively low expression in *D. radiodurans* and moderate expression in *E. coli*. Most applicable are those with high expression in *D. radiodurans* and low expression in *E. coli*, such as *PDR_1261* and *PDR_2508*, since these could serve for efficient protein expression in *D. radiodurans*. That is, these promoters could be expected to be inactive or show low activity while being shuttled though *E. coli* and high activity once incorporated into *D. radiodurans*.

Characterization of the top eight promoters. To gain a better understanding of how the selected putative promoters perform relative to promoters previously utilized to drive protein expression in *D. radiodurans*, we compared the strengths of the top candidates to nine previously published promoters selected from the literature (*PrecA, PamyE, PtufB, PkatA, PlexA, PmutL, PrecQ, PclpB*, and the positive-control, *PgroES*). Similar to the construction of the putative promoters described in Materials and Methods, we amplified promoter regions from the genome using target loci described in their original publications (Table 1), cloned them into the reporter construct, and assessed the fluorescence (Fig. 3A). Three of the benchmark promoters fell into the range of the top eight promoters selected from the initial screen. It is worth noting that two promoters, *PDR_1261* and *PrpmB*, were stronger than or comparable to strains harboring the four benchmark promoters exhibiting the highest fluorescent signals, *PgroES, PecA, PkatA*, and *PtufB*.

DR_1261 is currently annotated as a hypothetical protein and predicted by STRINGdb to have interactions with the *rsr* protein, which contributes to UV resistance by binding to small RNAs in *D. radiodurans* (28). The fluorescence driven by the promoter of DR_1261 was comparable to levels driven by PrecA and PkatA, proteins involved in stress response. If *PDR_1261* does interact with regulatory proteins, as predicted by STRINGdb, it could potentially play a regulatory role in *D. radiodurans*. Known stress response genes or ribosomal proteins were also annotated downstream of six of the top eight promoters and may help to maintain homeostasis of *D. radiodurans* under a wide variety of conditions, potentially explaining why they appeared to drive high expression under the conditions tested.

FIG 3 Characterization and benchmarking of the top eight promoters. (A) Normalized average fluorescence data for the top eight promoters and nine benchmarks previously used in the literature performed in exponential phase with biological triplicates. Error bars indicate standard deviations. Statistical significance is denoted by asterisks (**, $P < 0.01$), as determined by Student's t test. (B) Fluorescence microscopy images of promoter strains imaged at ×630 show cell-to-cell heterogeneity in fluorescence that varies across promoter strains. Scale bar, 20 μm. (C) Growth curves of top eight promoter strains, $PgroES$, and EV determined in triplicate. Cells were diluted to an OD_{600} of 0.1 and grown at 32°C with shaking in fresh TGY. The absorbance at 600 nm was measured hourly by a plate reader. (D) Normalized average fluorescence of top eight promoters, top three benchmarks, and $PgroES$ in $D.\ radiodurans$ in the exponential and stationary phases determined in biological triplicates. Error bars represent standard deviations. Statistical significance is denoted by asterisks (*, $P < 0.05$), as determined by Student's t test.

In tandem to the flow cytometry analysis, we visually inspected cultures for cell-to-cell variation in fluorescence by microscopy. This constitutes an important analysis since heterogeneous performance in engineered strains could result in suboptimal efficiency in biotechnological applications. Microscopy images of cells containing the top eight promoters, the top three benchmark promoters ($PrecA$, $PkatA$, and $PtufB$), and $PgroES$ are shown in Fig. 3B. While general fluorescence patterns observed via microscopy reflected the relative fluorescent trends found by flow cytometry, the heterogeneity seen by microscopy indicated a large range of cell-to-cell variation in expression driven by the selected promoters in $D.\ radiodurans$. The microscopy images for the strains carrying promoters PDR_1261 and $PrpmB$ showed that a majority of the cells were highly fluorescent, while the PDR_2508 and PDR_1473 strains showed only a few highly fluorescing cells, with the majority of cells displaying background levels of fluorescence. It is also worth noting that while the benchmark promoter sequences of $katA$, $recA$, and $groES$ drove high fluorescence, these strains also showed more cell-to-cell variability relative to constructs driven by PDR_1261 and $PrpmB$, the top promoters identified through this study. Similar trends in the cell-to-cell variation in expression driven by the promoters described above was also observed in saturated cultures (Fig. S1). Thus, these novel putative promoters are both strong and robust relative to other available promoters for $D.\ radiodurans$. The heterogeneity of these constructs also

highlights the difference between assessing promoter strength on an individual basis through cytometry compared to culture-based assessments of protein expression traditionally used, such as β-galactosidase assays, which may be influenced by strong outliers. Since these are native promoters, the biological mechanisms for the heterogeneity seen in this study may have involved complex interactions with native factors.

Native promoters used in their host organisms may retain the ability to interact with host systems, potentially interfering with the host's ability to perform normal metabolic processes. An impairment in metabolic function that influences cell fitness would greatly hinder the use of these promoters in industrial or biosensing applications. To investigate whether the top eight promoters in *D. radiodurans* would compromise cellular metabolism, changes in the growth rates of these strains were assessed (Fig. 3C). As seen in Fig. 3C, no significant differences were found in growth rates between cells expressing GFP from the top eight promoters and cells expressing GFP from the *groES* promoter, the promoter most commonly used in constitutive protein production in *D. radiodurans*.

It has been previously established that biological processes can be affected by key factors such as pH, temperature, acidity, and growth stage, which could be influenced by promoter activity levels (29). Therefore, we measured GFP fluorescence driven by *PgroES*, *PrecA*, *PkatA*, *PtufB*, and the top eight promoters in both exponential and stationary-phase culture conditions (Fig. 3D). For nearly all of the tested promoters, no significant difference was found between the fluorescence of exponential-phase and stationary-phase cultures. Encouragingly, *PDR_1261* and *PrpmB*, the two strongest promoters, were also the strongest regardless of condition, further indicating their potential use as strong, robust promoters that could be reliable in bioengineering applications. Cells containing the promoter regions from *DR_1348* and *DR_1261* were the only two to exhibit increased fluorescence in saturated cultures, demonstrating their potential use in applications utilizing saturated cultures such as the production of secondary metabolites. *PDR_1261* and *PDR_1348* are currently annotated as hypothetical proteins and therefore their potential regulatory effects and biological relevance in the native system remain uncharacterized.

Minimization of putative promoters. Although any novel promoter characterized for a nonmodel organism can be an enabling technology in and of itself, optimization through minimization can reduce promoter size for easier cloning (i.e., by Gibson assembly) and has resulted in the identification of important regulatory elements such as UTR elements, enhancer regions, or sigma factors (15, 29, 30). We used the two strongest promoter candidates, *PDR_1261* and *PrpmB*, a relatively small promoter with moderate activity, *PrplL*, and the smallest novel promoter, *PgroEL*, for minimization. Truncations from the 5′ end of the putative promoter region were generated by PCR, and cells harboring the truncated promoters were assessed for loss of fluorescence relative to those containing the full-length candidate (Fig. 4).

The promoter region of *DR_1261* was truncated in 10 bp increments, with a significant and abrupt decrease in activity observed between −20 and −30 bp, suggesting that this region is critical for transcription initiation; we were thus able to reduce the putative promoter size from 89 to 67 bp (Fig. 4A). The promoter region of *groEL* was truncated in 10 bp increments as well, but no significant decrease in fluorescence was observed, even after removal of 40 bp from the 74-bp full-length promoter, resulting in construction of a 34-bp promoter (Fig. 4B). Interestingly, removal of 40 bp actually resulted in a significant increase in promoter activity. An advantage of this minimal design is that the 34-bp *PgroEL* promoter is short enough to incorporate by Gibson assembly upstream of a gene of interest; therefore, we did not perform any further minimization on this promoter. The *PrplL* promoter was another one of the smallest original promoters that we investigated, with an initial size of 79 bp. To further minimize this promoter, we first truncated its sequence by 10-bp intervals until a total deletion of 60 bp was removed and a significant reduction in fluorescence was observed. The 40-bp truncation was unable to be constructed. This effort reduced the
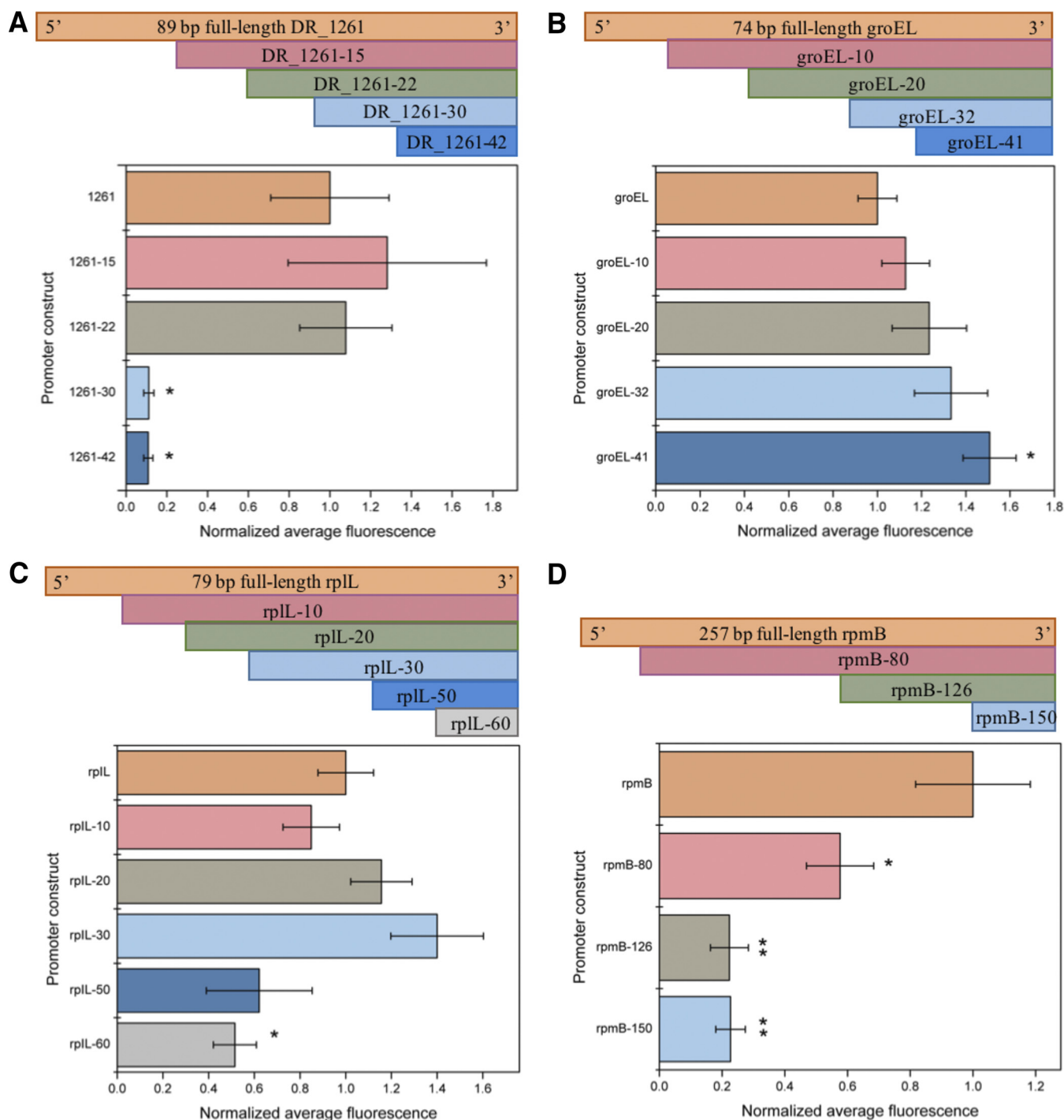
FIG 4 Normalized average fluorescence for promoter constructs where truncations were generated in regularly spaced intervals from the 5′ terminus until a loss in fluorescence was observed relative to the full-length promoter. Strains were assessed in exponential phase in triplicates, and error bars indicate the standard deviations. Statistical significance is denoted by asterisks (*, $P < 0.05$; **, $P < 0.01$), as determined by Student's t test. (A) Fluorescence data and schematic for the promoter of *DR_1261* and its truncations. (B) Fluorescence data and schematic for the promoter of *groEL* and its truncations. (C) Fluorescence data and schematic for the promoter of *rplL* and its truncations. (D) Fluorescence data and schematic for the promoter of *rpmB* and its truncations.

putative *rplL* promoter from 79 to 29 bp, resulting in an approximately 60% reduction in size (Fig. 4C). Lastly, the *rpmB* promoter was the longest of the four promoters we used for minimization, with an initial size of 257 bp (Fig. 4D). The effort resulted in a significant reduction in fluorescence between deletions of 80 and 120 bp, suggesting that an enhancing region or the transcription start site may be found within this region.

To identify whether the truncations influenced the cellular heterogeneity of protein production, fluorescence microscopy was performed on strains carrying *PgroES*, EV, and each minimized promoter and its full-length counterpart (Fig. S2). Similar to the assessment of the full-length promoters, the minimized promoters displayed a large range of cell-to-cell variation in fluorescence within cultures.

Computational assessment of characterized promoter regions of *D. radiodurans*. To identify potential bacterial promoter patterns that could be used as the basis for biological or evolutionary insight on *D. radiodurans* gene regulation, we assessed selected promoter strains using computational web tools. We first performed a BLAST analysis on the nucleotide sequence of the intergenic regions. Regions from the promoters of *DR_2508*, *groEL*, and *katA* showed similarity to genomic regions from *Deinococcus wulumugiensis*. Intergenic regions upstream of *tufB* and *DR_1473* contained tRNA sequences that showed conservation across several bacterial species, but regions outside the tRNA were unique to *D. radiodurans*. The remainder of the intergenic regions for the newly characterized promoters were unique to *D. radiodurans*. We then used several promoter prediction/analysis tools (BPROM, The Neural Network Promoter Predictor, and bTSSfinder) to investigate whether the selected promoters displayed any known promoter consensus regions. Several of the top candidate regions, such as *PrpmB* and *PtufB*, have predicted −10 and −35 sites by BPROM, but no promoter regions were predicted for *PDR_1261* or *PkatA* (Table S1). Complementing these results, the Neural Network Promoter Predictor identified four potential promoter sites, including two from *PdnaK* and one each from *PDR_2508* and *PrpmB* (Table S2). Lastly, we used bTSSfinder to identify potential sigma factor binding sites in the top eight candidates. The resulting output (Table S3) showed several predicted sigma factor binding sites in *PDR_2508*, two in *PDR_1473*, and one in *PrpmB*, indicating promoter activity. It is interesting that all of these tools identified *PrpmB*, which was one of the strongest promoters identified in this study, supporting the experimental results. While the prediction tools used in this study were able to identify some of the experimentally validated regions for promoter activity, none of the computational tools were able to accurately identify them all. Other promoters, such as *PDR_1261*, the strongest identified in the screen, were not predicted by any of the web tools. These results highlight the need for better genetic characterization in *D. radiodurans* and advances in computational tools to accurately make predictions in nonmodel organisms.

Conclusion. While the current suite of available promoters has been sufficient for single gene product expression, the introduction of pathways, such as those required for effective bioremediation of toxic compounds, into *D. radiodurans* will require an array of well-characterized promoters and regulatory elements for synthetic biology activities such as sensors, kill switches, and potentially tunable regulators (31–33). This study outlines the identification, screening, and characterization of a suite of novel promoters relative to a panel of previously published sequences for use in metabolic engineering applications with *D. radiodurans*. Through these characterization studies, two promoters—*PDR_1261* and *PrpmB*—displayed higher gene expression and homogeneity relative to any of the other promoters screened. Additional characterization revealed that gene expression driven by the majority of the putative promoter regions was relatively consistent across growth phase, making them more versatile for applications involving extended growth periods or transitions between growth phases. Four of the top eight promoters identified in this study were truncated to create minimal promoters and assessed via flow cytometry and microscopy. These minimized promoters of *PDR_1261*, *PrpmB*, *PgroEL*, and *PrplL* were at least 20% smaller than their full-length counterparts but still retained comparable levels of fluorescence and heterogeneity. Computational assessment of putative promoter regions demonstrated varied success in identifying promoter motifs among the characterized promoters, highlighting the need for experimental approaches and improvements in computational tools. Overall, this work expands the repertoire of promoters available for

D. radiodurans, resulting in a panel of robust promoters of various strengths and attributes that could be useful for a wide range of applications.

MATERIALS AND METHODS

Candidate selection. Candidate promoter sequences were identified through a bioinformatics analysis of previously published transcriptomics data (NCBI SRA accession no. GSM1584931: Sham_irradiated; *D. radiodurans* R1; ncRNA-Seq) (11). Transcriptomics data were collected in accordance with previously published protocols (34). Data were mapped to the genome sequence of *D. radiodurans* retrieved from the NCBI genome database (NC_001263) with Bowtie2 (35, 36). Read counts across each gene were then calculated with HT-Seq (37).

To identify potential promoter regions, candidate sequences were filtered by biological and technical parameters. Genes with read counts of <1% of the maximum number of reads per gene were removed from consideration for candidate selection to reduce the influence of low counts on candidate selection. In addition, the length of the upstream intergenic region was filtered to include only sequences between 45 and 500 bp. This lower limit was selected to reduce the number of false positives due to multiple genes encoded on operons, since the majority of operon intergenic regions are less than 40 bp in length in *E. coli*, and bioinformatics analyses suggest that most are under 50 bp in other bacterial species (38–40). The upper limit of intergenic length was set to 500 bp to identify putative promoters with sizes amenable for cloning into a plasmid. Transcripts per million (TPM) were calculated for each gene to normalize the number of read counts by the downstream gene length. The resulting data set was sorted by TPM and candidate promoter sequences were selected across a range of predicted strengths for experimental assessment (see Data Set S1 in the supplemental material).

***Post hoc* assessment of promoter candidates with selected promoter prediction software.** After primary assessment and selection of the top 25% (eight) candidates, the intergenic regions containing putative promoters were assessed by prediction software to search for features potentially indicative of efficient gene expression in *D. radiodurans*. The motif-finding tool, MEME, and promoter prediction software such as BPROM (Softberry), bTSSfinder (*E. coli* type setting), and Neural Network Promoter Prediction (prokaryote setting, minimum promoter score 0.8) were used to search for motifs indicative of promoter regions within the selected candidate promoters (41–43). Putative promoter regions were also assessed by BLAST to search for conservation across other taxa.

Strain development. The plasmid pRadGroGFP, which has been previously used to assess gene expression levels, was cut with restriction sites XhoI and BamHI (17). A gblock (IDT) was then used to replace the ApaI site with a SacII site. The pRadGro construct was chosen for promoter screening in *D. radiodurans* because it has been previously characterized to have a plasmid copy number approximately equal to the genome copy number (44). The resulting plasmid was either digested with XhoI and SacII or PCR amplified with primers pRadBB-F and pRadBB-R (listed in Table 1), digested with DpnI at 37°C for 1 h, purified with a Zippy DNA clean and concentrate kit, and then digested with XhoI and SacII at 37°C for 2 h. The backbone was then gel extracted for cloning of candidate promoter regions using an Illustra GFX PCR DNA and gel band purification kit (GE Healthcare).

Inserts were PCR amplified with primers (Table 1) designed to target intergenic regions between two *cis*-oriented adjacent genes identified by the bioinformatics analysis. These primers were used to amplify candidate promoter regions from *D. radiodurans* R1 genomic DNA template. Amplification by these primers incorporated a 5' XhoI site, the first three amino acids of groES (Met-Leu-Lys), and SacII site on the 3' end. Addition of the first three amino acids was found to improve the signal of the reporter protein. The PCR products were cut with XhoI and SacII at 37°C for 2 h and column purified. T4 DNA ligase (NEB) was used to perform 10-μl ligations with 100 ng of DNA at insert/vector ratios of 3:1 or 6:1, followed by incubation at 16°C overnight. The following day, ligations were desalted and transformed into DH5α or DH10β electrocompetent cells (NEB) and recovered with 1 ml of super optimal broth. Cells were recovered at 37°C for 1 h with shaking, then plated on Luria-Bertani (LB) agar plates containing 50 μg/ml carbenicillin and grown overnight at 37°C. Colonies were assessed by colony PCR using *Taq* DNA polymerase (Invitrogen) and GFP colony forward and reverse primers (Table 1). Colonies resulting in successful amplification were grown overnight in LB medium. Plasmids were isolated according to the manufacturer's instructions (Zippy Miniprep kit) and then submitted for Sanger sequence verification at the University of Texas at Austin Genome Sequencing and Analysis Facility (GSAF) with the reverse GFP colony PCR primer. Once confirmed, 1 μg of plasmid DNA was chemically transformed into *D. radiodurans*, as described previously (45). Cells were plated on tryptone-glucose-yeast extract (TGY) agar containing 0.34 μg/ml chloramphenicol at 32°C for 3 days to allow the growth of transformed colonies.

Fluorescence assessment. After 3 days of growth, single colonies of transformed *D. radiodurans* were inoculated in triplicate into liquid TGY containing 0.34 μg/ml chloramphenicol and grown overnight at 32°C. The following day, cells were diluted 1:20 and grown to an optical density at 600 nm (OD$_{600}$) of 0.3 to 0.8. Next, 50 μl of culture was transferred into 1 ml of 1$\times$ PBS and assessed for fluorescence by cytometry using a FACSCalibur or FACSAria Fusion (Becton Dickinson) with a 15-mW 488-nm argon-ion laser and a 515- to 545-nm emission filter. The top eight candidates from the preliminary screen were selected for further characterization via cytometry relative to empty vector (EV) control and alongside a suite of previously published promoters to enable standardization against promoters currently available for use in *D. radiodurans* (12–15, 19, 46–48).

Influence of growth phase on promoter activity. Transformed *D. radiodurans* harboring promoter candidates selected from the preliminary screen were inoculated in triplicate into liquid TGY containing 0.34 μg/ml chloramphenicol and grown overnight at 32°C. The following day, saturated cultures were diluted 1:20 and grown to an OD$_{600}$ of 0.3 to 0.8. Then, 50 μl of either saturated or exponential culture

was transferred into 1 ml of 1× PBS and assessed for fluorescence by cytometry using a FACSCalibur at 488 nm.

Sequence verification of transformed *D. radiodurans*. To validate that the selected *D. radiodurans* were successfully transformed with the plasmid, a 5-ml culture of each candidate strain was grown overnight at 32°C. The following day, the cell pellet was isolated by centrifugation (10 min at 4,000 × *g*) and resuspended in 1 ml of nuclease-free H_2O. Cells were lysed via mini-Beadbeater (100 s × 2; Biospec Products), and cell components were removed by centrifugation at 13,000 × *g* for 2 min. Then, 600 μl of the supernatant was processed to retrieve plasmids from *D. radiodurans* according to the Zippy miniprep protocol. Plasmids were then transformed into *E. coli* strain DH5α and plated on LB containing 50 μg/ml carbenicillin. Transformants were cultured and grown overnight in LB medium at 37°C. The following day, the plasmids were isolated by miniprep, and the sequence was confirmed by the GSAF.

Microscopy. Microscopy was performed on selected *D. radiodurans* strains containing putative promoter-bearing plasmids from the primary screen. A single colony of transformed cells was inoculated into 5 ml of TGY liquid containing 0.34 μg/ml chloramphenicol. The following day, the culture was diluted 1:20 with TGY containing chloramphenicol and grown at 32°C to an OD_{600} of 0.4 to 0.6. Cells were isolated from 500 μl of cell culture by centrifugation (2 min at 5,000 × *g*) and resuspended in 200 μl of 1× PBS for assessment by fluorescence microscopy (Zeiss Axiovert 200M fluorescence microscope) at the University of Texas at Austin's Microscopy and Imaging Facility, with exposure times of 340 ms for fluorescent images and 34 ms for bright-field imaging with a 63× lens objective. The same procedure and analysis were also repeated for saturated cultures.

Growth curve analysis. The promoters considered in this study are native to *D. radiodurans* and could influence cellular metabolism due to interactions with endogenous regulatory systems. To assess the impact of the promoters on the growth rate, growth curves were generated for the selected strains to ensure that introduction of the putative promoter did not impact the growth rate of the strain. To construct the growth curve, single colonies of transformed *D. radiodurans* cells were inoculated in triplicate into 5 ml of TGY liquid containing 0.34 μg/ml chloramphenicol and grown overnight at 32°C. The following day, the cultures were diluted 1:20 in 5 ml with TGY containing chloramphenicol and grown for 2 h at 32°C. The OD_{600} was measured and cultures were diluted to an OD_{600} of 0.1 in 20 ml. After dilution, cultures were grown at 32°C with shaking, and 200-μl aliquots were used to measure the OD_{600} hourly with a plate reader (Bio-Tek) until the stationary phase was reached.

Promoter minimization. Promoter regions of four genes were chosen for minimization because of their sizes and/or strengths: *DR_1261*, *rpmB*, *rplL*, and *groEL*. Primers were designed to systematically reduce the size of the promoter PCR product from the 5′ terminus (Table 1). Cloning was performed as described previously with cut sites XhoI and SacII into the pRadGroGFP construct containing SacII. In the case of *rpmB*, gblocks (IDT) were used rather than PCR products due to difficulties in PCR amplification, so *rpmB* was assessed with truncations of 80, 120, and 150 bp. gblocks containing the promoter region and GFP were cloned using Gibson Assembly into the backbone generated by amplification of pRadGroGFP with pRad-BBforGA forward and reverse primers (Table 1). Fluorescence was assessed by cytometry, as described above, to determine the influence of the deletion on promoter activity. *D. radiodurans* cultures containing the minimized promoters were then assessed by microscopy to determine whether minimization influenced expression heterogeneity.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/AEM.01356-19.

SUPPLEMENTAL FILE 1, PDF file, 0.8 MB.

REFERENCES

1. Daly MJ. 2000. Engineering radiation-resistant bacteria for environmental biotechnology. Curr Opin Biotechnol 11:280–285. https://doi.org/10.1016/S0958-1669(00)00096-3.
2. Appukuttan D, Rao AS, Apte SK. 2006. Engineering of *Deinococcus radiodurans* R1 for bioprecipitation of uranium from dilute nuclear waste. Appl Environ Microbiol 72:7873–7878. https://doi.org/10.1128/AEM.01362-06.
3. Gerber E, Bernard R, Castang S, Chabot N, Coze F, Dreux-Zigha A, Hauser E, Hivin P, Joseph P, Lazarelli C, Letellier G, Olive J, Leonetti JP. 2015. *Deinococcus* as new chassis for industrial biotechnology: biology, physiology, and tools. J Appl Microbiol 119:1–10. https://doi.org/10.1111/jam.12808.
4. Brim H, McFarlan SC, Fredrickson JK, Minton KW, Zhai M, Wackett LP, Daly MJ. 2000. Engineering *Deinococcus radiodurans* for metal remedia-

tion in radioactive mixed waste environments. Nat Biotechnol 18:85–90. https://doi.org/10.1038/71986.

5. Lange CC, Wackett LP, Minton KW, Daly MJ. 1998. Engineering a recombinant *Deinococcus radiodurans* for organopollutant degradation in radioactive mixed waste environments. Nat Biotechnol 16:929–933. https://doi.org/10.1038/nbt1098-929.

6. Misra CS, Mukhopadhyaya R, Apte SK. 2014. Harnessing a radiation inducible promoter of *Deinococcus radiodurans* for enhanced precipitation of uranium. J Biotechnol 189:88–93. https://doi.org/10.1016/j.jbiotec.2014.09.013.

7. Leistra AN, Curtis NC, Contreras LM. 2019. Regulatory noncoding sRNAs in bacterial metabolic pathway engineering. Metab Eng 52:190–214. https://doi.org/10.1016/j.ymben.2018.11.013.

8. Marques CR. 2018. Extremophilic microfactories: applications in metal and radionuclide bioremediation. Front Microbiol 9:1–10.

9. Masters CI, Smith MD, Gutman PD, Minton KW. 1991. Heterozygosity and instability of amplified chromosomal insertions in the radioresistant bacterium *Deinococcus radiodurans*. J Bacteriol 173:6110–6117. https://doi.org/10.1128/jb.173.19.6110-6117.1991.

10. Telang S, Patel P, Sarangdhar V, Donde S. 2014. Isolation and cloning of the endoglucanase gene from *Bacillus pumilus* and its expression in *Deinococcus radiodurans*. Biotech 4:57–65. https://doi.org/10.1007/s13205-013-0127-3.

11. Tsai C-H, Liao R, Chou B, Contreras LM. 2015. Transcriptional analysis of *Deinococcus radiodurans* reveals novel small RNAs that are differentially expressed under ionizing radiation. Appl Environ Microbiol 81:1754–1764. https://doi.org/10.1128/AEM.03709-14.

12. Lecointe F, Coste G, Sommer S, Bailone A. 2004. Vectors for regulated gene expression in the radioresistant bacterium *Deinococcus radiodurans*. Gene 336:25–35. https://doi.org/10.1016/j.gene.2004.04.006.

13. Meima R, Rothfuss HM, Gewin L, Lidstrom ME. 2001. Promoter cloning in the radioresistant bacterium *Deinococcus radiodurans*. J Bacteriol 183:3169–3175. https://doi.org/10.1128/JB.183.10.3169-3175.2001.

14. Kota S, Dhamodharan V, Pradeepkumar PI, Misra HS. 2015. G-quadruplex forming structural motifs in the genome of *Deinococcus radiodurans* and their regulatory roles in promoter functions. Appl Microbiol Biotechnol 99:9761–9769. https://doi.org/10.1007/s00253-015-6808-6.

15. Schmid AK, Lidstrom ME. 2002. Involvement of two putative alternative sigma factors in stress response of the radioresistant bacterium *Deinococcus radiodurans*. J Bacteriol 184:6182–6189. https://doi.org/10.1128/jb.184.22.6182-6189.2002.

16. Holland AD, Rothfuss HM, Lidstrom ME. 2006. Development of a defined medium supporting rapid growth for *Deinococcus radiodurans* and analysis of metabolic capacities. Appl Microbiol Biotechnol 72:1074–1082. https://doi.org/10.1007/s00253-006-0399-1.

17. Villa JK, Amador P, Janovsky J, Bhuyan A, Saldanha R, Lamkin TJ, Contreras LM. 2017. A genome-wide search for ionizing- radiation-responsive elements in *Deinococcus radiodurans* reveals a regulatory role for the DNA gyrase subunit A gene's 5′ untranslated region. Appl Environ Microbiol 83:1–19.

18. Cheng J, Wang H, Xu X, Wang L, Tian B, Hua Y. 2015. Characteristics of dr1790 disruptant and its functional analysis in *Deinococcus radiodurans*. Braz J Microbiol 46:601–611. https://doi.org/10.1590/S1517-838246220131018.

19. Dulermo R, Onodera T, Passot F. 2015. Identification of new genes contributing to the extreme radioresistance of *Deinococcus radiodurans* using a Tn*5*-based transposon mutant library. PLoS One 10:e0124358. https://doi.org/10.1371/journal.pone.0124358.

20. Wang J, Julin DA. 2004. DNA helicase activity of the RecD protein from *Deinococcus radiodurans*. J Biol Chem 279:52024–52032. https://doi.org/10.1074/jbc.M408645200.

21. Jolivet E, Lecointe F, Coste G, Satoh K, Narumi I, Bailone A, Sommer S. 2006. Limited concentration of RecA delays DNA double- strand break repair in *Deinococcus radiodurans* R1. Mol Microbiol 59:338–349. https://doi.org/10.1111/j.1365-2958.2005.04946.x.

22. Bentchikou E, Servant P, Coste G, Sommer S. 2007. Additive effects of SbcCD and PolX deficiencies in the *in vivo* repair of DNA double-strand breaks in *Deinococcus radiodurans*. J Bacteriol 189:4784–4790. https://doi.org/10.1128/JB.00452-07.

23. Passot FM, Nguyen HH, Dard-Dascot C, Thermes C, Servant P, Espéli O, Sommer S. 2015. Nucleoid organization in the radioresistant bacterium *Deinococcus radiodurans*. Mol Microbiol 97:759–774. https://doi.org/10.1111/mmi.13064.

24. de la Tour CB, Blanchard L, Dulermo R, Ludanyi M, Devigne A, Armengaud J, Sommer S, de Groot A. 2015. The abundant and essential HU proteins in *Deinococcus deserti* and *Deinococcus radiodurans* are translated from leaderless mRNA. Microbiology 161:2410–2422. https://doi.org/10.1099/mic.0.000186.

25. Yu T, Zhou YJ, Huang M, Liu Q, Pereira R, David F, Nielsen J. 2018. Reprogramming yeast metabolism from alcoholic fermentation to lipogenesis. Cell 174:1549–1558.e14. https://doi.org/10.1016/j.cell.2018.07.013.

26. Yang Y, Shen W, Huang J, Li R, Xiao Y, Wei H, Chou Y-C, Zhang M, Himmel ME, Chen S, Yi L, Ma L, Yang S. 2019. Prediction and characterization of promoters and ribosomal binding sites of *Zymomonas mobilis* in system biology era. Biotechnol Biofuels 12:1–13. https://doi.org/10.1186/s13068-019-1399-6.

27. Li S, Wang J, Li X, Yin S, Wang W, Yang K. 2015. Genome-wide identification and evaluation of constitutive promoters in streptomycetes. Microb Cell Fact 14:1–11. https://doi.org/10.1186/s12934-015-0351-0.

28. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, Von Mering C. 2015. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43:D447–D452. https://doi.org/10.1093/nar/gku1003.

29. Madsen SM, Arnau J, Vrang A, Givskov M, Israelsen H. 1999. Molecular characterization of the pH-inducible and growth phase-dependent promoter P170 of *Lactococcus lactis*. Mol Microbiol 32:75–87. https://doi.org/10.1046/j.1365-2958.1999.01326.x.

30. Singh P, Chachan S, Singhi D, Srivastava P. 2016. Isolation and molecular characterization of a stationary phase promoter useful for gene expression in *Gordonia*. Gene 591:153–160. https://doi.org/10.1016/j.gene.2016.07.018.

31. Gao GJ, Fan L, Lu HM, Hua YJ. 2008. Engineering *Deinococcus radiodurans* into biosensor to monitor radioactivity and genotoxicity in environment. Sci Bull 53:1675–1681. https://doi.org/10.1007/s11434-008-0224-6.

32. Vazquez-Anderson J, Contreras LM. 2013. Charming gene management styles for synthetic biology applications. RNA Biol 10:1778–1797. https://doi.org/10.4161/rna.27102.

33. Rothfuss HM. 2004. Genetic engineering of *Deinococcus radiodurans* R1 for bioremediation of mixed waste. PhD dissertation. University of Washington, Seattle, WA.

34. Gelderman G, Contreras LM. 2013. Discovery of posttranscriptional regulatory RNAs using next generation sequencing technologies, p 269–295. *In* Alper HS (ed), Systems metabolic engineering: methods and protocols. Humana Press, Totowa, NJ.

35. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.

36. White O, Eisen JA, Heidelberg JF, Hickey EK, Peterson JD, Dodson RJ, Haft DH, Gwinn ML, Nelson WC, Richardson DL, Moffat KS, Qin H, Jiang L, Pamphile W, Crosby M, Shen M, Vamathevan JJ, Lam P, McDonald L, Utterback T, Zalewski C, Makarova KS, Aravind L, Daly MJ, Minton KW, Fleischmann RD, Ketchum KA, Nelson KE, Salzberg S, Smith HO, Venter JC, Fraser CM. 1999. Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. Science 286:1571–1577. https://doi.org/10.1126/science.286.5444.1571.

37. Anders S, Pyl PT, Huber W. 2015. HTSeq: a Python framework to work with high-throughput sequencing data. Bioinformatics 31:166–169. https://doi.org/10.1093/bioinformatics/btu638.

38. Tjaden B, Haynor DR, Stolyar S, Rosenow C, Kolker E. 2002. Identifying operons and untranslated regions of transcripts using *Escherichia coli* RNA expression analysis. Bioinformatics 18:S337–S344. https://doi.org/10.1093/bioinformatics/18.suppl_1.S337.

39. Price MN, Huang KH, Alm EJ, Arkin AP. 2005. A novel method for accurate operon predictions in all sequenced prokaryotes. Nucleic Acids Res 33:880–892. https://doi.org/10.1093/nar/gki232.

40. Salgado H, Moreno-Hagelsieb G, Smith TF, Collado-Vides J. 2000. Operons in *Escherichia coli*: genomic analyses and predictions. Proc Natl Acad Sci U S A 97:6652–6657. https://doi.org/10.1073/pnas.110147297.

41. Reese MG. 2001. Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. Comput Chem 26:51–56. https://doi.org/10.1016/S0097-8485(01)00099-7.

42. Solovyev V, Salamov A. 2010. Automatic annotation of microbial genomes and metagenomic sequences. Metagenomics Appl Agric 2010:1–18.

43. Shahmuradov IA, Mohamad Razali R, Bougouffa S, Radovanovic A, Bajic VB. 2017. bTSSfinder: a novel tool for the prediction of promoters in cyanobacteria and *Escherichia coli*. Bioinformatics 33:334–340.

44. Meima R, Lidstrom M. 2000. Characterization of the minimal replicon of a

cryptic *Deinococcus radiodurans* SARK plasmid and development of versatile *Escherichia coli-D. radiodurans* shuttle vectors. Appl Environ Microbiol 66: 3856–3867. https://doi.org/10.1128/AEM.66.9.3856-3867.2000.

45. Smith MD, Lennon E, McNeil LB, Minton KW. 1988. Duplication insertion of drug resistance determinants in the radioresistant bacterium *Deinococcus radiodurans*. J Bacteriol 170:2126–2135. https://doi.org/10.1128/jb.170.5.2126-2135.1988.

46. Turner EH, Lauterbach K, Pugsley HR, Palmer VR, Dovichi NJ. 2007. Detection of green fluorescent protein in a single bacterium by capillary electrophoresis with laser-induced fluorescence. Anal Chem 79:778–781. https://doi.org/10.1021/ac061778r.

47. Masters C, Minton K. 1992. Promoter probe and shuttle plasmids for *Deinococcus radiodurans*. Plasmid 28:258–261. https://doi.org/10.1016/0147-619X(92)90057-H.

48. Funayama T, Narumi I, Kikuchi M, Kitayama S, Watanabe H, Yamamoto K. 1999. Identification and disruption analysis of the *recN* gene in the extremely radioresistant bacterium *Deinococcus radiodurans*. Mutat Res 435:151–161. https://doi.org/10.1016/s0921-8777(99)00044-0.