

## Supplementary Information

### Development of a highly sensitive whole-cell biosensor for arsenite detection through engineered promoter modifications

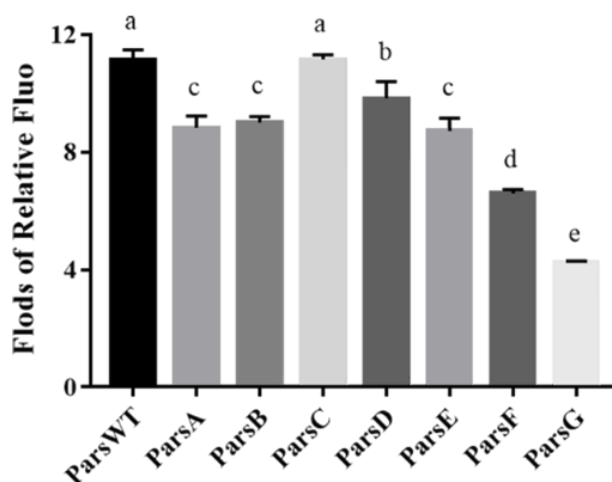
Sheng-Yan Chen<sup>a</sup>, Wenping Wei<sup>b</sup>, Bin-Cheng Yin<sup>b</sup>, Yanbin Tong<sup>a</sup>, Jianjiang Lu<sup>a</sup>, Bang-Ce Ye<sup>a,b,c\*</sup>

<sup>a</sup>*School of Chemistry and Chemical Engineering, Shihezi University, Shihezi, 832003, China.*

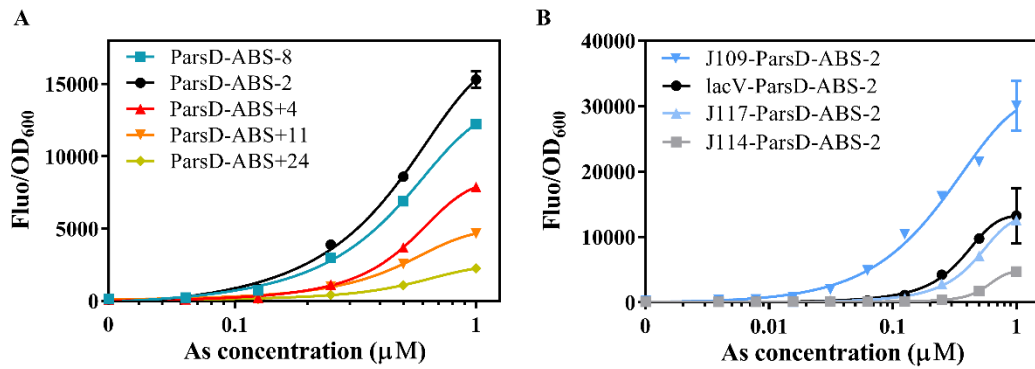
<sup>b</sup>*Lab of Biosystem and Microanalysis, State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, Shanghai, 200237, China.*

<sup>c</sup>*Institute of Engineering Biology and Health, Collaborative Innovation Center of Yangtze River Delta Region Green Pharmaceuticals, College of Pharmaceutical Sciences, Zhejiang University of Technology, Hangzhou 310014, Zhejiang, China*

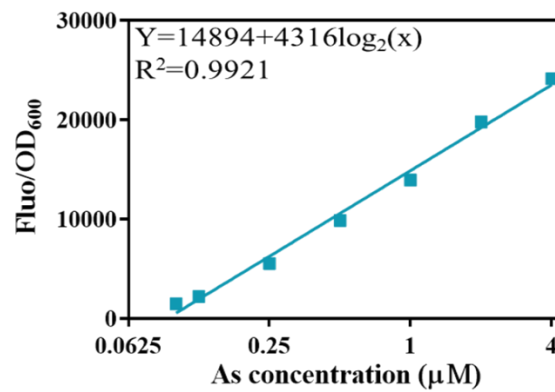
#### Supplementary Figures



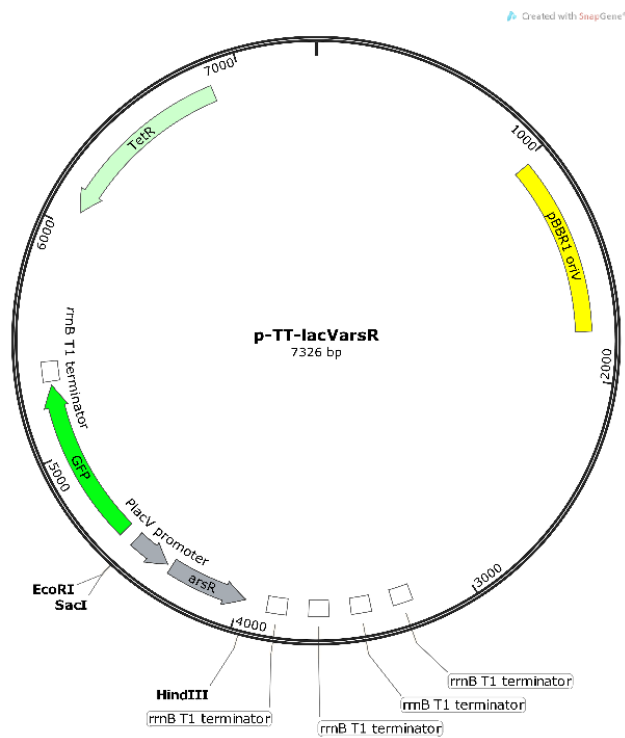
**Figure S1.** Data from Figure 2A, but expressed as the ratio between the Fluo/OD<sub>600</sub> in the arsenite-containing sample and that of sample without arsenite. Means values  $\pm$  SD (n=3). Means in the column followed by the different letter differ significantly at  $p \leq 0.05$ .



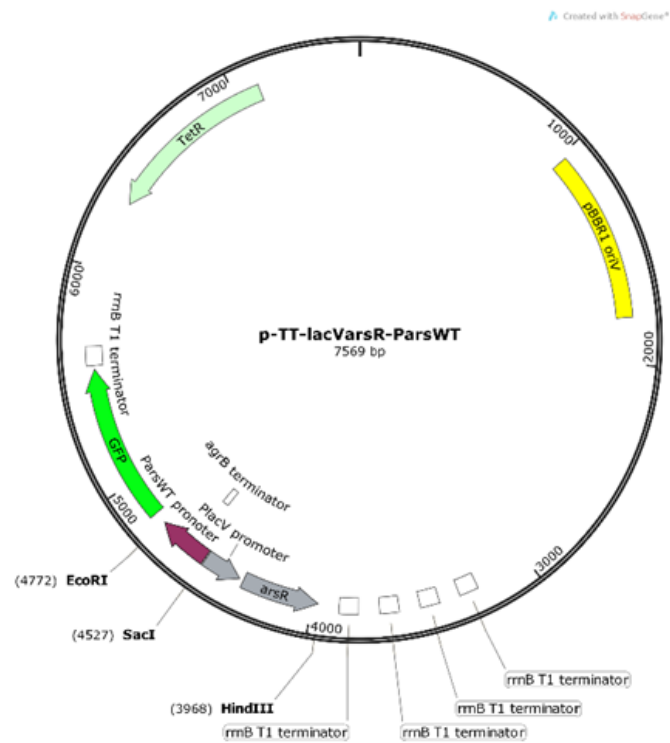
**Figure S2. The location of the second ArsR binding site and the expression level of ArsR were adjusted to further improve the sensor performance.** (A) The fluorescence level of the sensor in response to arsenite concentrations when the second ArsR binding site is placed at a different position relative to the promoter transcription initiation site. (B) The fluorescence level of the sensor in response to arsenite concentrations when different activity promoters (J109, J117, J114, the promoter activity gradually increased from left to right.) to drive the expression of ArsR. Error bars are smaller than the data symbols (n=3).



**Figure S3.** Calibration curve of arsenite with concentration range from 0.1 to 4 μM.  $R^2$  values and nonlinear regression equation result from linear interpolation of the averages from three replicates, average fluorescence of each sample was achieved after induction for 4 h at 25 °C. Error bars are smaller than the data symbols.



**Figure S4.** Schematic illustration of the p-TT-lacVarsR vector.



**Figure S5.** Schematic illustration of the p-TT-lacVarsR-ParsWT vector.

## Supplementary Tables

**Table S1.** The effect of the addition of a second binding site on different active promoters.

Report variant	Add As	Output <sup>a</sup>	Output ratio <sup>b</sup>
ParsWT	1 $\mu$ M	7839	100%
ParsWT-ABS+4		1198	15.3% (1198/7839)
ParsD		25795	100%
ParsD-ABS+4		8209	31.8% (8209/25795)
ParsG		43886	100%
ParsG-ABS+4		34695	79% (34695/43886)

<sup>a</sup>The output values are averages of Fluo/OD<sub>600</sub> from triplicate measurements (added 1  $\mu$ M arsenite incubation for 4 h at 30 °C). <sup>b</sup> The fluorescence output ratio of different active promoters containing single ABS was defined as 100%, and the fluorescence output ratio of different active promoters containing double ABS was defined as output ratio=(dual ABS output)/(single ABS output).

**Table S2.** The plasmids used in this study.

Plasmid name	Arsenite regulated promoter sequence
ParsWT	: GCACTTACACATTCGTTAAGTCATATATGTTTT <b>GACTT</b> ATCCGCTT-CGAAGAGA-// <b>GACACT</b> ACCTGCAACA <b>GAATTC</b>
ParsA	: //TATGTTACCTGCAACA <b>GAATTC</b>
ParsB	: //TACTGTACCTGCAACA <b>GAATTC</b>
ParsC	: //TAAATTACCTGCAACA <b>GAATTC</b>
ParsD	: //GATACTACCTGCAACA <b>GAATTC</b>
ParsE	: //GATAATACCTGCAACA <b>GAATTC</b>
ParsF	: //TATAGTACCTGCAACA <b>GAATTC</b>
ParsG	: //TATAATACCTGCAACA <b>GAATTC</b>
ParsWT-ABS+4	: //GACACTACCTGCAACAATACACATTCGTTAAGTCATATATGTTTT <b>GACTT</b> ATCCGC <b>GAATTC</b>
ParsD-ABS+4	: //GATACTACCTGCAACAATACACATTCGTTAAGTCATATATGTTTT <b>GACTT</b> ATCCGC <b>GAATTC</b>
ParsD-ABS-2	: //GATACTACCTGCTACACATTCGTTAAGTCATATATGTTTT <b>GACTT</b> ATCCGC <b>GAATTC</b>
ParsD-ABS+11	: //GATACTACCTGCAACA <b>GAATTC</b> ACTATTACTACACATTCGTTAAGTCATATATGTTTT <b>GACTT</b> ATCCGC <b>GAATTC</b>
ParsD-ABS+24	: //GATACTACCTGCAACA <b>GAATTC</b> ACTATTACCTTCCTCTGCACTTACACATTCGTTAAGTCATATATGTTTT <b>GACTT</b> ATCCGC <b>GAATTC</b>
ParsG-ABS+4	: //TATAATACCTGCAACAATACACATTCGTTAAGTCATATATGTTTT <b>GACTT</b> ATCCGC <b>GAATTC</b>
ParsD-ABS-8	: //GATACTTACACATTCGTTAAGTCATATATGTTTT <b>GACTT</b> ATCCGC <b>GAATTC</b>
ParsG-ABS-8	: //TATAATTACACATTCGTTAAGTCATATATGTTTT <b>GACTT</b> ATCCGC <b>GAATTC</b>
Plamisd	The sequence of four different promoters and RBS that drive the expression of arsR.
lacV-ParsD-ABS-2	( <i>SacI</i> )ACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCT <b>TTACGTTG</b> TGAGCGCTCACAATTATAGTGTGTGGAATCaatcaggagcgcaat(arsR)
J114-ParsD-ABS-2	( <i>SacI</i> ) <b>TTTACAGCTAGCTCAGTCCTAGG</b> ACTGTGCTAGCtctagagattaaagagagagaaatactag(arsR)
J117-ParsD-ABS-2	( <i>SacI</i> ) <b>TTGACAGCTAGCTCAGTCCTAGG</b> ATTGTGCTAGCtctagagattaaagagagagaaatactag(arsR)
J109-ParsD-ABS-2	( <i>SacI</i> ) <b>TTTACAGCTAGCTCAGTCCTAGG</b> ACTGTGCTAGCtctagagattaaagagagagaaatactag(arsR)

Note: The sequence difference of all Pars promoter variants is only between the double slash and the *EcoRI* restriction site (GAATTC). All plasmid except ParsD-ABS-2 used the lacV promoter to drive expression of ArsR, ParsD-ABS-2 uses four different promoters (lacV, J114, J117, and J109) to drive expression of ArsR. The bold sequences represent the -10 or -35 site. The ArsR recognition site (a proximal ABS) with underline. Transcript start sites and *EcoRI* restriction site are colored red.

**Table S3.** The primers used in this study.

Primer name	Sequence
HA-arsR-F	gctcacaattatagtggtggaatCAATCAGGAGCGCAATATGTCATTTC
HR- <i>Hind</i> III-arsR-R	ggaattggggatcggaagcttTTAACTGCAAATGTTCTTACTGTCC
HR- <i>Sac</i> I-lacV-F	gaataagccgaattcgagctcGCGCCCAATACGCAAACC
HA-lacV-R	attccacacactataattgtgagcGCTCACAACGTAAAGCCTGGGGTGCCTAA TGAGTG
HR- <i>Sac</i> I-Pars-F	cggtttgcgtattgggcgcgagctcAGCCACTGGCTAATAGTATTGAGCT
HR- <i>Eco</i> RI-Pars-R	gtagttaggggaataagccgaattcTTGTTGCAGGTAGTGTCTCTCTTCG
Pars-AR	agttaggggaataagccgaattcTTGTTGCAGGT <b>AACATA</b> TCTCTTCGAAGCG GATAAGT
Pars-BR	agttaggggaataagccgaattcTTGTTGCAGGT <b>ACAGTA</b> TCTCTTCGAAGCG GATAAGT
Pars-CR	agttaggggaataagccgaattcTTGTTGCAGGT <b>AATTTA</b> TCTCTTCGAAGCG GATAAGT
Pars-DR	agttaggggaataagccgaattcTTGTTGCAGGT <b>AGTATC</b> TCTCTTCGAAGCG GATAAGT
Pars-ER	agttaggggaataagccgaattcTTGTTGCAGGT <b>ATTATC</b> TCTCTTCGAAGCG GATAAGT
Pars-FR	agttaggggaataagccgaattcTTGTTGCAGGT <b>ACTATA</b> TCTCTTCGAAGCG GATAAGT
Pars-GR	agttaggggaataagccgaattcTTGTTGCAGGT <b>ATTATA</b> TCTCTTCGAAGCG GATAAGT
Parswt-ABS+4-R	gtagttaggggaataagccgaattcGCGGATAAGTCAAAAACATATATGACTT AACGAATGTGTAGCGCTCCTGATTGTTGCAG
ParsD-ABS-2-R	gtagttaggggaataagccgaattcGCGGATAAGTCAAAAACATATATGACTT AACGAATGTGTAGCAGGTAGTATCTCTCTTCGAAGC
ParsD-ABS+4-R	gtagttaggggaataagccgaattcGCGGATAAGTCAAAAACATATATGACTT AACGAATGTGTATTGTTGCAGGTAGTATCTCTCTTCG
ParsD-ABS+11-R	gtagttaggggaataagccgaattcGCGGATAAGTCAAAAACATATATGACTT AACGAATGTGTAGTAATAGTTGTTGCAGGTAGTATCTCTCTTC
ParsD-ABS+24-R	gtagttaggggaataagccgaattcGCGGATAAGTCAAAAACATATATGACTT AACGAATGTGTAAGTGCAGAGGAAGGTAATAGTTGTTG
ParsG-ABS+4-R	gtagttaggggaataagccgaattcGCGGATAAGTCAAAAACATATATGACTT AACGAATGTGTATTGTTGCAGGTATTATATCTCTTCG
ParsD-ABS-8-R	ttagttaggggaataagccgaattcGCGGATAAGTCAAAAACATATATGACTT AACGAATGTGTAAGTATCTCTCTTCGAAGCGGATAAGTC
ParsG-ABS-8-R	ttagttaggggaataagccgaattcGCGGATAAGTCAAAAACATATATGACTT AACGAATGTGTAATTATATCTCTTCGAAGCGGATAAGTC
ParsD-ABS-XX-R	AGTGCAGAGGAAGGTAATAGTTGTTGCAGGTAGTATCTCTCTT CGAAG
J-xx-R	TGCTAGCTCTAGAGATTAAAGAGGAGAAATACTAGATGTCATT TCTGTTACCCATCCAAT

J114-R	atactattagccagtggct <u>gagctc</u> TTTATGGCTAGCTCAGTCCTAGGTACAA TGCTAGCTCTAGAGATTAAAGAGGAG
J117-R	atactattagccagtggct <u>gagctc</u> TTGACAGCTAGCTCAGTCCTAGGGATTG TGCTAGCTCTAGAGATTAAAGAGGAG
J109-R	atactattagccagtggct <u>gagctc</u> TTTACAGCTAGCTCAGTCCTAGGGACTG TGCTAGCTCTAGAGATTAAAGAGGAG
pPROBE-TT-F	GGAATTGGGGATCGGAAGCTT (sequencing primer)
pPROBE-TT-R	GCATCACCTTCACCCTCTCCAC (sequencing primer)

Note: The DNA sequence (lowercase) indicates the homologous recombination sequence on the plasmid. The underlined sequences are restriction endonuclease sites.

**Table S4.** Sequences of the genes.

Gene name	Sequence and characteristic
lacV promoter-RBS <sub>arsR</sub>	( <i>SacI</i> )ACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGG CACGACAGGTTTCCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAAT GTGAGTTAGCTCACTCATTAGGCACCCCAGGCT <b>TTTACG</b> TTGTGAGCGCTCAC AAT <b>TATAGT</b> GTGTGGAATCAATCAGGAGCGCAAT (arsR) -35site: <b>TTTACG</b> -10site: <b>TATAGT</b>
arsR	(lacV) <b>ATG</b> TCATTTCTGTTACCCATCCAATTGTTCAAAATTCTTGCTGATGAAACCCGT CTGGGCATCGTTTTACTGCTCAGCGAACTGGGAGAGTTATGCGTCTGCGATCT CTGCACTGCTCTCGACCAGTCGCAGCCCAAGATCTCCCGCCACCTGGCATTG CTGCGTGAAAGCGGGCTATTGCTGGACCGCAAGCAAGGTAAGTGGGTTCATT ACCGCTTATCACCGCATATTCCAGCATGGGCGGCGAAAATTATTGATGAGGCC TGGCGATGTGAACAGGAAAAGTTTCAGGCGATTGTCCGCAACCTGGCTCGA CAAAACTGTTCCGGGGACAGTAAGAACATTTGCAGT <b>TAA</b> ( <i>HindIII</i> )
ParsWT promoter	( <i>SacI</i> )AGTATTGAGCTGTTAGATAAGAACTCTCTCACTCCAGCCAGAGCCACCA ACTCAGGGCTGGAAAGT <b>AAAAAACCGACGCAAAGTCGGTTTTTTTACGTC</b> <b>CTGATT</b> CAGACCTCCTTTCAAATGAATAGCCAACTCAAATTCACACCTATTA CCTTCCTCTGCACT <b>TACACATTCGTTAAGTCATATATGTTTTGACTTA</b> TCCG CTTCGAAGAGAG <b>ACACT</b> ACCTGCAACAA ( <i>EcoRI</i> ) agrB Terminator: <b>AAAAAACCGACGCAAAGTCGGTTTTTTTACGTCCTGA</b> (Weel-Sneve et al. 2013) ArsR Binding site: <b><u>TACACATTCGTTAAGTCATATATGTTTTGACTTA</u></b> -35site: <b>TTGACT</b> -10site: <b>GACACT</b>
GFP	( <i>EcoRI</i> )GGCTTATTCCCTAACTAACTAAAGATTAACCTTTATAAGGAGGAAAAAC AT <b>ATG</b> AGTAAAGGAGAAGAAGCTTTTCACTGGAGTTGTCCCAATTCTTGTTGA ATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGAGAGGGTGAA GGTGATGCAACATACGGAAAGCTTACCCTTAAATTTATTTGCACTACTGGAAA ACTACCTGTTCCATGGCCAACACTTGTCACTACTTTGACTTATGGTGTTCAAT GCTTTTCAAGATACCCAGATCATATGAAACGGCATGACTTTTTTCAAGAGTGCC ATGCCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTTCAAAGATGACGGGA ACTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCTTGTTAATAG AATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGACAC AAATTGGAATACAACATACTCACACAATGTATACATCATGGCAGACAAACA AAAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGA AGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCC TGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAG ATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGC TGGGATTACACATGGCATGGATGAACTATACAAAT <b>AG</b>

Note: Parentheses before and after the gene are directly linked genes or cleavage sites, and the features of promoter are indicated. Start codon and stop codon are marked in red.



## Reference

Weel-Sneve, R., Kristiansen, K.I., Odsbu, I., Dalhus, B., Booth, J., Rognes, T., Skarstad, K., Bjoras, M., 2013. Single transmembrane peptide DinQ modulates membrane-dependent activities. ***PLoS genetics*** 9(2), e1003260.