Supplementary Materials for

High-throughput transcriptional characterization of regulatory sequences from bacterial biosynthetic gene clusters

Jimin Park[1,2], Sung Sun Yim[1,3], Harris H. Wang[1,3*]

Affiliations

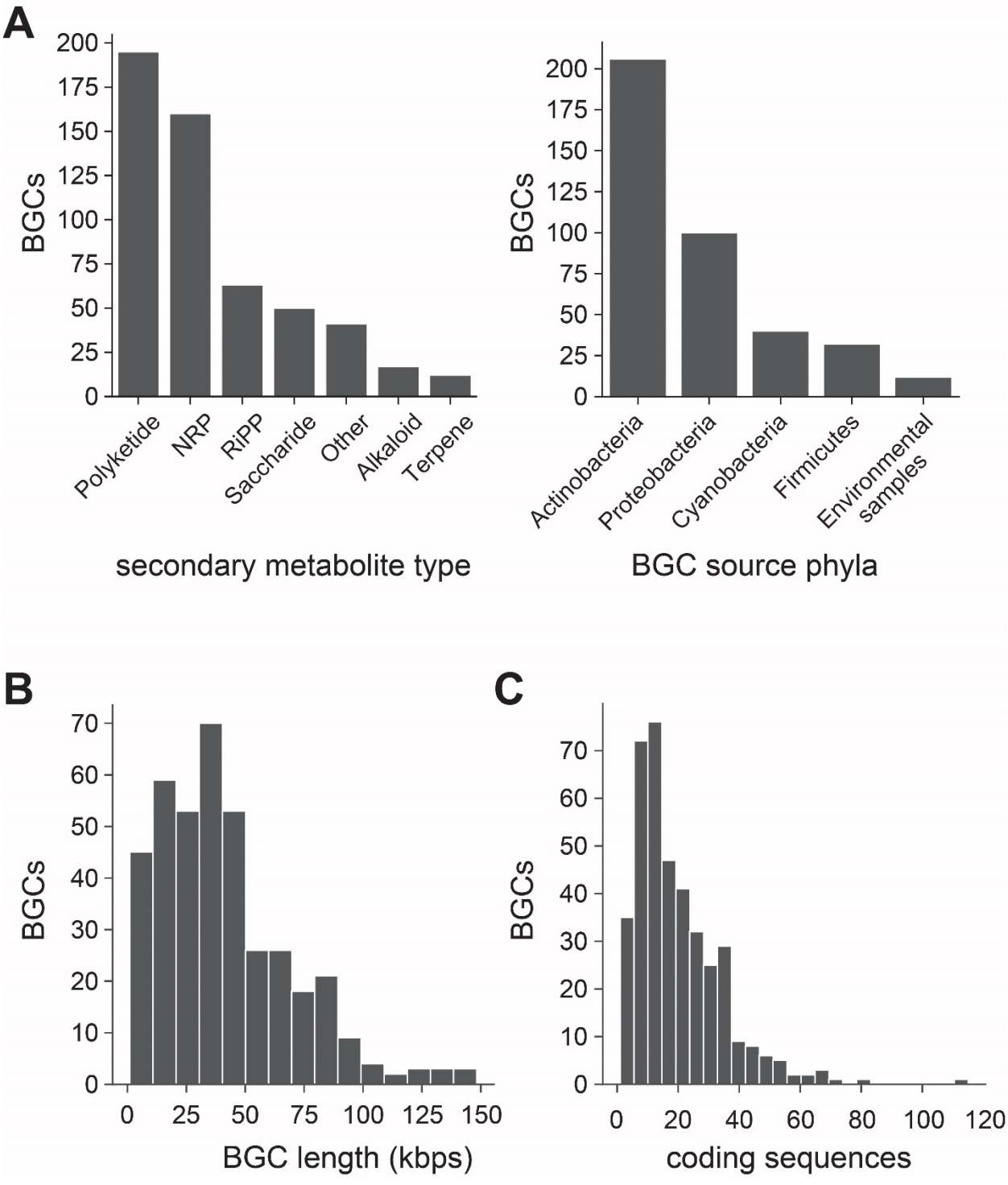
[1] Department of Systems Biology, Columbia University Irving Medical Center, New York, NY, 10032, USA.

[2] Integrated Program in Cellular, Molecular and Biomedical Studies, Columbia University Irving Medical Center, New York, NY, 10032 USA

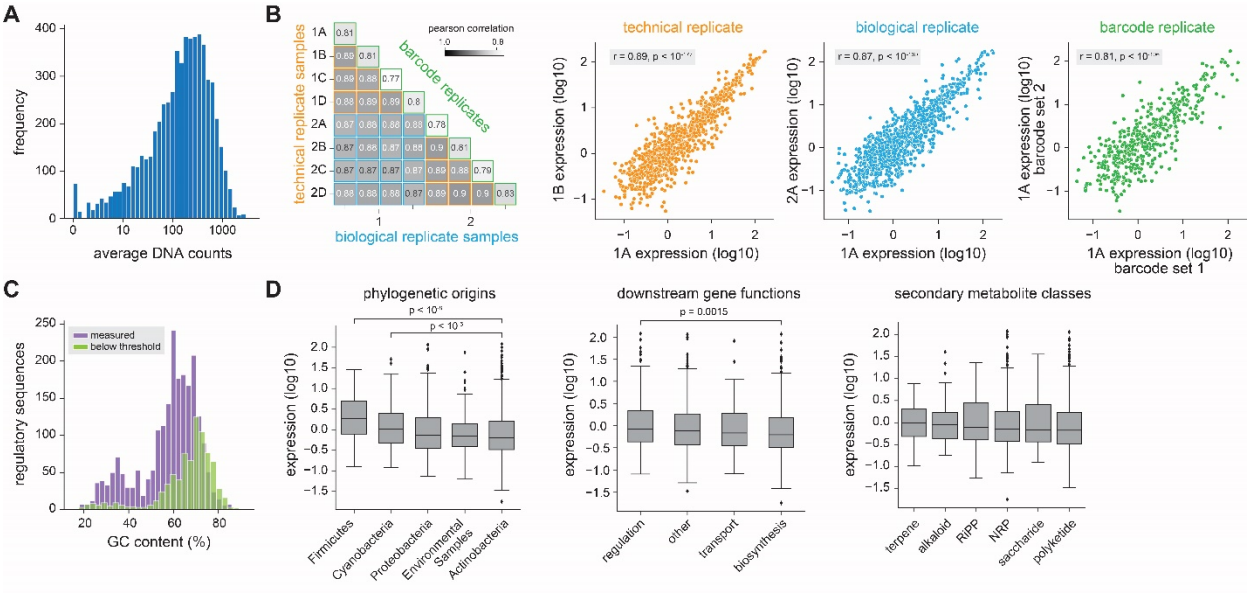
[3] Department of Pathology and Cell Biology, Columbia University Irving Medical Center, New York, NY, 10032, USA

* Correspondence to: hw2429@columbia.edu

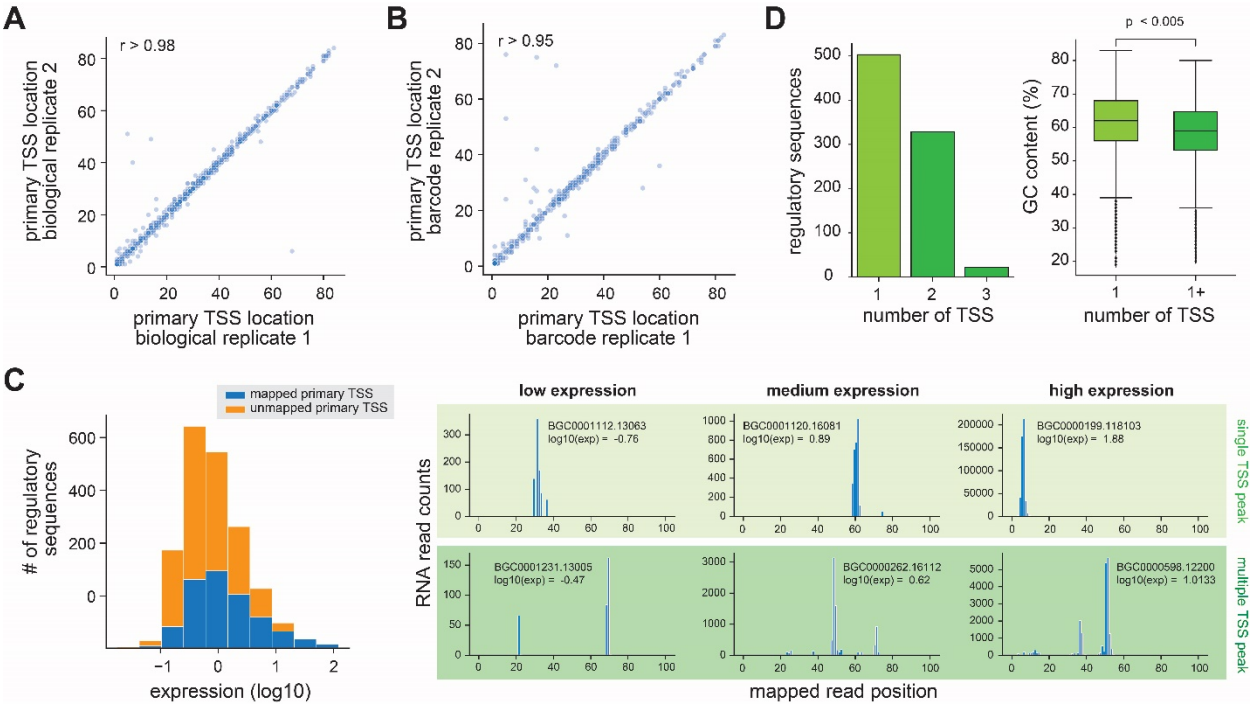
Supplementary Figures



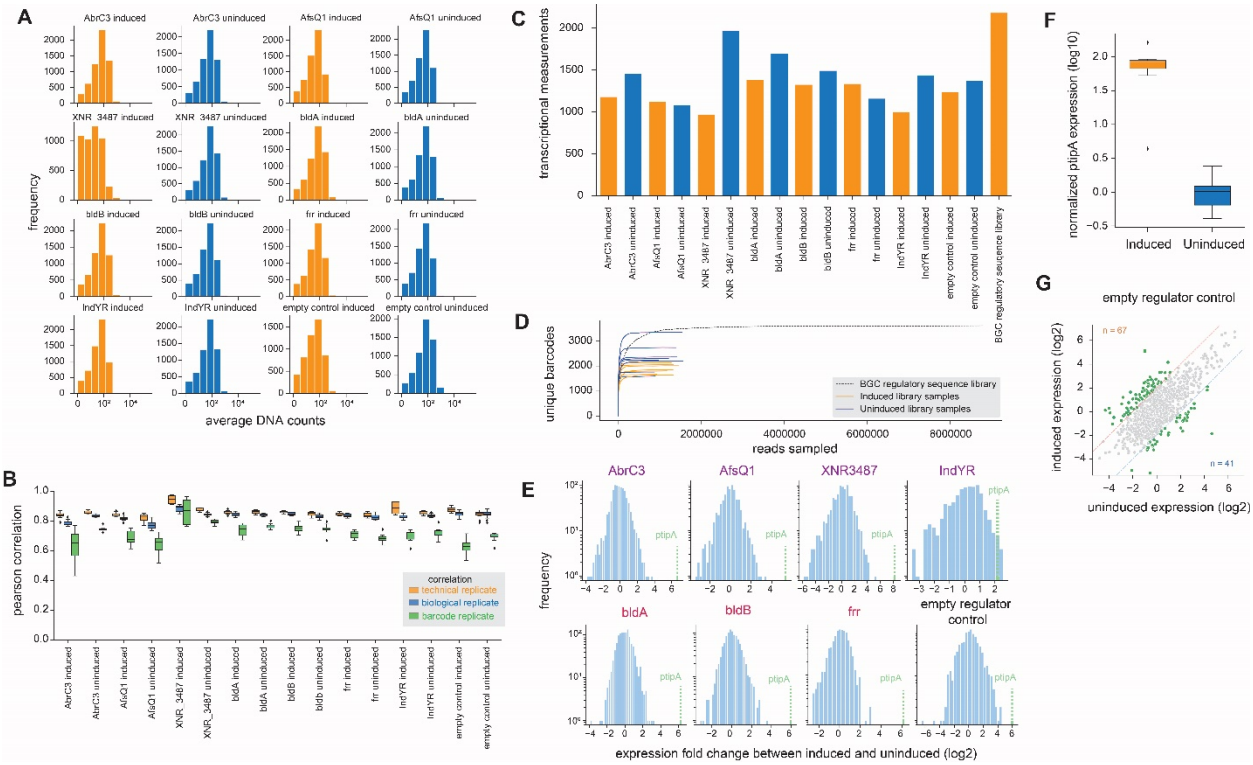
Supplemental Figure S1. (A) Bargraphs displaying Biosynthetic Gene Cluster counts by their secondary metabolite type categories (left). BGCs counts by their phylogenetic source (right). **(B)** Distribution of BGC lengths in kilobases. **(C)** Distribution of number of coding sequences per BGC.



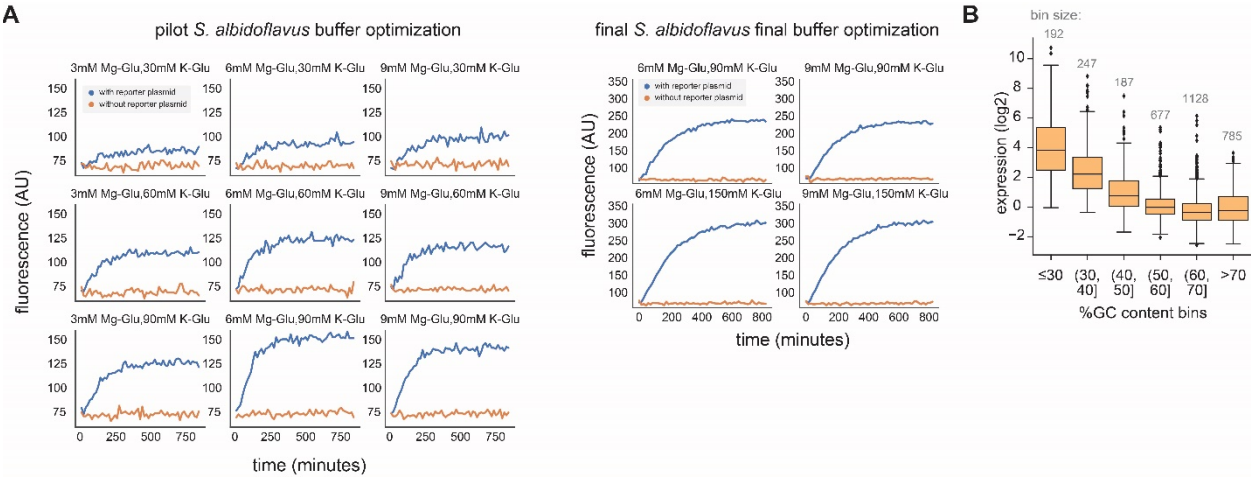
Supplemental Figure S2. (A) Distribution of regulatory sequence library in *S. albidoflavus* as measured by average DNA barcode abundances. **(B)** Correlation between technical, biological, and barcode replicate measurements of transcriptional levels of BGC regulatory sequences. Heatmap shows all correlations and the scatter plot shows a single representative correlation for technical (orange), biological (blue), and barcode (green) replicate comparisons. **(C)** GC content distribution of regulatory sequences with measured transcriptional levels (blue) and sequences with transcriptional levels below detection threshold (orange). **(D)** Transcriptional levels of the BGC regulatory sequence library stratified by phylogenetic origins (left), downstream gene functions (middle), and secondary metabolite classes (right). Distributions that were significantly different, as determined using a t-test, are denoted with the respective p-values.



Supplemental Figure S3. (A) Primary TSS locations between two biological replicates. **(B)** Primary TSS locations between identical regulatory sequences measured with two different barcodes. **(C)** Expression distribution of regulatory sequences with mapped primary TSS (blue) and unmapped primary TSS (orange). Below, random selection of 6 regulatory sequences with varying expression (high/mid/low) and number of TSS peaks (single/multiple) and their mapped read positions are illustrated. **(D)** (Left) Number of mapped TSS locations for regulatory sequences that have a primary TSS. (Right) Comparison of GC content distribution between sequences from the left plot that have a single TSS against sequences that have multiple TSS.



Supplemental Figure 4. (A) Relative DNA abundances of regulatory sequences in each regulator-regulatory sequence sample. **(B)** Distribution of technical, biological, and barcode correlation displayed as boxplots for each regulator-regulatory sequence library sample. **(C)** Number of regulatory sequences with transcriptional measurements across each regulator-regulatory sequence library samples. **(D)** Rarefaction curves illustrating the number of unique RNA barcodes recovered from increasing number of reads sampled from various RS and regulator-RS library samples. (**E**) Histogram of expression fold change between induced and uninduced samples of each regulator-RS library samples with *ptipA* expression fold change denoted by green dashed line. **(F)** Boxplot of normalized *ptipA* expression levels across all library samples. **(G)** Normalized expression of regulatory sequences from the empty regulator control samples compared between induced and uninduced conditions.



Supplementary Figure 5. **(A)** Optimization of Mg-Glutamate and K-glutamate concentrations for *S. albidoflavus* cell free expression system using the RNA aptamer broccoli and DHFBI-1T as a reporter. Left plot shows a preliminary large-scale optimization while the right plot shows a final optimization comparing fewer reaction conditions. **(B)** Box plot showing transcriptional activity of RS measured *in vitro* grouped by their GC content. Center line denote the median, with boxes showing top and bottom 25th percentiles. Whiskers are 1.5x the respective boxes and values beyond the whiskers are displayed as individual points.

Supplementary Data

Supplementary Data 1. Regulatory sequence libraries and metadata

Supplementary Data 2. BGC Library expression data in *S. albidoflavus*

Supplementary Data 3. BGC Library expression data with global regulatory proteins

Supplementary Data 4. Defined regulatory sequence library

Supplementary Data 5. BGC library expression data with *S. albidoflavus* cell free system

Supplementary Tables

Supplementary Table 1. Global Regulatory factors and their effects on BGC regulatory sequence expression

Regulator	**Description**	**DNA binding?**	**# of upregulated regulatory sequences**	**# of downregulated regulatory sequences**
AbrC3	Two component system (TCS) Response Regulator	Yes	16	24
AfsQ1	TCS Response Regulator	Yes	13	19
XNR_3847	TCS Response Regulator	Yes	65	38
IndYR	GntR-like transcriptional regulator	Yes	10	77
bldA	TTA codon tRNA	No	12	6
bldB	Pleiotropic small protein	No	15	8
frr	Ribosome recycling factor	No	5	15

Supplementary Table 2. List of plasmids used in this study

Plasmid	Description	Source/Reference
pIJ10257	Streptomyces shuttle vector with ermE promoter, Hyg^R, and *Φ*BT1 integration system	[1]
pJP12	pIJ10257 plasmid containing eGFP reporter gene	This work
pJP50	pIJ10257 derivative, replaced ermE expression loci with insulated terminators, ATG-less mCherry reporter, and a library cloning site.	This work
pIJ6902	Streptomyces shuttle vector with thiostrepton inducible promoter, tsr (thiostrepton resistance gene), AAC(3)-IV (apramycin resistance gene) and *Φ*BT1 integration system	[2]
pIJ6902-abrC3	pIJ6902 containing thiostrepton inducible abrC3 regulator gene	This work
pIJ6902-afsQ1	pIJ6902 containing thiostrepton inducible afsS regulator	This work
pIJ6902-XNR-3487	pIJ6902 containing thiostrepton inducible XNR-3487 regulator	This work
pIJ6902-bldA	pIJ6902 containing thiostrepton inducible bldA regulator	This work
pIJ6902-bldB	pIJ6902 containing thiostrepton inducible bldB regulator	This work
pIJ6902-frr	pIJ6902 containing thiostrepton inducible frr regulator	This work
pIJ6902-lndYR	pIJ6902 containing thiostrepton inducible lndYR regulator	This work
pJP19	pIJ6902 containing thiostrepton inducible mCherry reporter	This work

Supplementary Table 3. List of regulators used in this study

Regulator	Desciption	Sequence
abrC3	Two component system Response Regulator from *S. albus*	GTGCGGGTGGTCATCGCCGAGGATTCGGTGCTGCTGA GGGAGGGCCTGACCCGGCTCCTGACCGACCGGGGGC ACGAGGTGGTCGCCGGGGTCGGGGATGCCCAGGCGC TGGTGAAGACCATCGCCGAGCTGGCGGCGGCGGACGA GCTGCCGGATGTGGTGGTCGCGGACGTGCGGATGCCG CCGACCCACACCGACGAGGGCGTACGGGCTGCCGTGC AACTGCGGAGCAGGCATCCGGGACTCGGAGTGCTGGT GCTCTCCCAGTACGTGGAAGAGCGGTACGCGACGGAG CTGCTGGCCGGTTCCAGCAGGGGGGTCGGCTACCTGC TCAAGGACCGGGTGGCGGAGGTCCGCGAGTTCGTGGA CGCGGTGGTGCGGGTGGCCGAGGGCGGCACGGCCCT GGACCCGGAGGTGGTGGCCCAGCTCCTGGGCCGCAG CCGGAAGCAGGACGTGCTCGCCGCCCTGACCCCGCG GGAGCGCGAGGTGCTGGGCCTGATGGCCGAGGGCAG GACCAACTCGGCGATCGCCCGGCAGCTGGTGGTCAGC GATGGTGCGGTGGAGAAGCACGTGAGCAACATCTTCCT GAAGCTGGGCCTCTCGCCGAGCGACGGCGACCACCGC CGGGTCCTCGCCGTGCTGACGTACCTCAACTCCTGA
afsQ1	Two component system Response Regulator from *S. albidoflavus*	ATGGCGGGCGTGCCTTCCCTGTTGCTGATCGAGGACG ACGACGCCATCCGTACCGCCCTGGAGCTGTCGCTGAC ACGCCAGGGGCATCGGGTGGCGACCGCTGCCACCGG CGAGGAGGGTCTGACGCTCCTGCGCGAGCAGCGGCC GGACCTGATCGTGCTGGACGTGATGCTGCCGGGCATC GACGGCTTCGAGGTGTGCCGGCGCATCCGCCGCACGG ACCAGTTGCCGATCATCCTGCTGACCGCGCGCAGCGA CGACATCGACGTGGTGGTGGGGCTGGAGTCCGGCGCC GACGACTACGTGGTCAAGCCGGTCCAGGGCCGGGTGC TGGACGCGCGTATCCGGGCCGTGCTGCGGCGGGGCG AGCGGGAGTCGAACGACGCGGCGGTCTTCGGGTCGCT GGTGATCGACCGGGCCGCGATGACGGTCACGAAGAAC GGCGAGGACCTGCAACTGACCCCGACCGAGCTGCGGT TGCTCCTGGAGCTGTCCCGCAGGCCGGGGCAGGCGCT CTCCCGGCAGCAGTTGCTGCGGCTGGTCTGGGAGCAC GACTACCTCGGTGACTCGCGGCTGGTCGACGCCTGTG TGCAGCGGCTGCGCGCCAAGGTGGAGGACGTGCCGTC CTCGCCGACGCTGATCCGTACCGTGCGCGGCGTGGGC TACCGGCTGGACGCGCCTCAGTGA
XNR-3487	Two component system Response Regulator from *S. albidoflavus*	ATGGCCGACCACACCCATGTCCTGTTCGTGGAGGACG ACGACGTGATCCGGGAGGCGACCCAGCTCGCCCTGGA GCGGGTCGGCTTCCGGGTGACCGCCATGCCCGACGG CCTCTCGGGGCTGGAGGCGTTCCGCGCCGACCGGCC GGACATCGCGCTGCTCGACGTGATGGTGCCCGGCATG GACGGGGTCAGCCTCTGCCGCCGCATACGCGACGAGT CGACCGTGCCGGTGATCATGCTCTCGGCGCGGGCCGA CTCCATCGACGTGGTCCTCGGGCTGGAGGCGGGCGCC GACGACTACGTCACCAAGCCGTTCGACGGCTCGGTCC TGGTCGCCCGCATCCGGGCCGTGCTGCGCCGCTTCGG CCACGCCGCCGGGGCGCAGGGCGCCAGAGCCGAGGA

		GGAGCCCTCCGGCGGGCCCGCCGAGGCCGACGGGCT GCTCCGCTTCGGCGATCTGGAGCTGGACACCGAGGGC ATGGTGGTGCGGCGGGCCGGCAGCCCGGTCGCGCTG ACGCCGACCGAGATGCGGCTGCTGCTGGAGTTCTCGG CCGCGCCGGGCACCGTCCTCTCCCGCGACAAGCTGCT GGAACGCGTCTGGGACTACGGCTGGGGCGGCGACAC CCGGGTGGTCGACGTGCATGTGCAGCGGCTGCGCGCC AAGATCGGCCAGGAGCGGATAGAGACCGTCCGCGGCT TCGGCTACAAACTGAGGGCCTGA
bldA	TTA codon tRNA from *S. albidoflavus*	GCCCGGATGGTGGAATGCAGACACGGCGAGCTTAAAC CTCGCTGCCCTTCACGGGCGTACCGGTTCGAGTCCGG TTCCGGGCA
bldB	Pleiotropic small protein from *S. coelicolor*	ATGGCCCAGGTGCCGGACGAGGACGTCAAAGCCCGCA AGGAGCGCGAGCGGGACGAGCTGTACGCGCTCGACAT CTCGGGTGTGGAGTGGCACAGCGCGCCGGGGACGGA GGAACACGAGGAGCGGGTGGAGATCGCCTATCTGCCC GACGGAGCCGTGGCCATGCGGTCGTCGCTGGATCCGG AGACGGTGCTGCGGTACACCGAGGCGGAGTGGCGGG CTTTCGTCCTGGGTGCGCGGGACGGGGAGTTCGACCT GGAGCCGGCGCCGGGCGACGGGGGCGTCGTCGCCGA GTGA
frr	Ribosome recycling factor from *S. albidoflavus*	GTGATCGAAGAGACCCTCCTCGAGGCCGAGGAGAAGA TGGAAAAGGCCGTCGTGGTCGCCAAGGAGGACTTCGC CGCGATCCGCACCGGCCGTGCGCACCCGGCGATGTTC AACAAGATCGTGGCGGACTACTACGGTGCCCTCACGC CGATCAATCAGCTGGCCTCGTTCTCCGTGCCGGAGCC GCGCATGGCCGTTGTCACGCCGTTCGACTCGAGCGCG CTGCGCAACATCGAGCAGGCCATCCGCGACTCGGACC TCGGCGTCAACCCGAGCAACGACGGCCGCATCATCCG CGTGTCGTTCCCGGAGCTGACCGAGGAGCGGCGCCGC GAGTTCATCAAGGTCGCCAAGGGCAAGGCCGAGGACT CGAAGATCTCGATCCGCGCCGTGCGCCGCAAGGCCAA GGAGACCATCGACAAGCTCGTCAAGGACGGCGAGGTC GGTGAGGACGAGGGCCGCCGCGGCGAGAAGGAACTC GACGACACCACCGCCAAGTACGTGGCTCAGGTGGACG AGCTGCTCAAGCACAAGGAAGCCGAGCTGCTCGAAGT CTGA
lndR	GntR-like transcriptional regulator from *S. globisporus*	GTGTCGGCAATCGAGCGGAAGGTGAAGTCCGTGGTCG TTTTCCGCATCGACCGGCGCAGTGGGGTGGCGACCTA TCTCCAGATCGTCCGGCAGGTCGAGCAGGCGCTGCGC ATGGGCGCGCTGGAGGAGGGCGACCGACTGCCCACG GCGGCCCAGGTCGCCGCGACCACGAAGGTCAACCCCA ACACGACCCTCAAGGCCTACCGCGAGCTGGAACGCAT GGGCCTCGCGGAGGTACGCCAGGGAGCGGGCACGTT CATCACCCGCACCCTCGCCCAGCCCCAGTCCGGCCCC GACTCCCCGCTCCGCACCGCCCTCACCGACTGGCTGA CCCGGGCCCGCGCCGAGGGCCTCAGGGGCCAGGACG TCACGGCGCTGTTCCACGCGGCGTTCGAGAACGCGTA CCCGGGCGAGGCGCCGGACTGA

Supplementary Table 4. List of oligonucleotides used in this study

Name	Sequence (5' to 3')
JP194	GCCTAGGTAACGTGGAAGT
JP195	GTCGAGTTGATCCATGGTC
JP262	GCCCCATATGATGGTGAGCAAGGGCGAGG
JP263	ATCACTCGAGTTACTTGTACAGCTCGTCCATG
JP349	GAATTCCCCAGATCTAAAGTTTTGTCG
JP350	CATATGTCCGCTCCCTTCTC
JP351	TCATTTACTAACGTCTGGAAAGACGACAAAACTTTAGATCTGGGGAATTCTCACTTGTACAGCTCGTCC
JP352	ACGTCACGTGAGGAGGCAGCGTGGACGGCGTCAGAGAAGGGAGCGGACATATGGTGTCCAAGGGCGAGG
JP616	TTCTTGGCCTTGTACGTCGTCTTGA
JP_adapter	/5Phos/NNATGTACTCTGCGTTGATACCACTGCTT/3SpC3/

Sequencing library prep oligonucleotides
For amp1, four primers of varying lengths were pooled and used together.
For amp2 primers, p5 and p7 indices are indicated in **bold**.

Name	Sequence (5' to 3')
DNA_lib_amp1_3N_F	CCTACACGACGCTCTTCCGATCTNNNCCTTGGACACCTGCAG
DNA_lib_amp1_4N_F	CCTACACGACGCTCTTCCGATCTNNNNCCTTGGACACCTGCAG
DNA_lib_amp1_5N_F	CCTACACGACGCTCTTCCGATCTNNNNDCCTTGGACACCTGCAG
DNA_lib_amp1_6N_F	CCTACACGACGCTCTTCCGATCTNNNNDRCCTTGGACACCTGCAG
DNA_lib_amp1_0N_R	GAGTTCAGACGTGTGCTCTTCCGATCTGGTACCCGGGGATCC
DNA_lib_amp1_1N_R	GAGTTCAGACGTGTGCTCTTCCGATCTTGGTACCCGGGGATCC
DNA_lib_amp1_2N_R	GAGTTCAGACGTGTGCTCTTCCGATCTCHGGTACCCGGGGATCC
DNA_lib_amp1_3N_R	GAGTTCAGACGTGTGCTCTTCCGATCTAHMGGTACCCGGGGATCC
RNA_lib_amp1_3N_F	CCTACACGACGCTCTTCCGATCTNNNCCTCCTCGCCCTTGGACAC
RNA_lib_amp1_4N_F	CCTACACGACGCTCTTCCGATCTNNNNCCTCCTCGCCCTTGGACAC
RNA_lib_amp1_5N_F	CCTACACGACGCTCTTCCGATCTNNNNNCCTCCTCGCCCTTGGACAC
RNA_lib_amp1_6N_F	CCTACACGACGCTCTTCCGATCTNNNNNNCCTCCTCGCCCTTGGACAC
RNA_lib_amp1_0N_R	GAGTTCAGACGTGTGCTCTTCCGATCTNNNAAGCAGTGGTATCAACGCAGAGTACAT

RNA_lib_amp1_1N_R	GAGTTCAGACGTGTGCTCTTCCGATCTNNNNAAGCAGTGGTATCAACGCAGAGTACAT
RNA_lib_amp1_2N_R	GAGTTCAGACGTGTGCTCTTCCGATCTNNNNNAAGCAGTGGTATCAACGCAGAGTACAT
RNA_lib_amp1_3N_R	GAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNAAGCAGTGGTATCAACGCAGAGTACAT
p5_bc1	AATGATACGGCGACCACCGAGATCTACAC**CTCTCTAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc2	AATGATACGGCGACCACCGAGATCTACAC**TATCCTCT**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc3	AATGATACGGCGACCACCGAGATCTACAC**AGAGTAGA**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc4	AATGATACGGCGACCACCGAGATCTACAC**GTAAGGAG**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc5	AATGATACGGCGACCACCGAGATCTACAC**ACTGCATA**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc6	AATGATACGGCGACCACCGAGATCTACAC**AAGGAGTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc7	AATGATACGGCGACCACCGAGATCTACAC**CTAAGCCT**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc8	AATGATACGGCGACCACCGAGATCTACAC**GCGTAAGA**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc9	AATGATACGGCGACCACCGAGATCTACAC**TAGATCGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc10	AATGATACGGCGACCACCGAGATCTACAC**CGTCTAAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc11	AATGATACGGCGACCACCGAGATCTACAC**TCTCTCCG**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc12	AATGATACGGCGACCACCGAGATCTACAC**TCGACTAG**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc13	AATGATACGGCGACCACCGAGATCTACAC**TTCTAGCT**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc14	AATGATACGGCGACCACCGAGATCTACAC**CCTAGAGT**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc15	AATGATACGGCGACCACCGAGATCTACAC**CTATTAAG**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p5_bc16	AATGATACGGCGACCACCGAGATCTACAC**AAGGCTAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
p7_bc1	CAAGCAGAAGACGGCATACGAGAT**TAAGGCGA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc2	CAAGCAGAAGACGGCATACGAGAT**CGTACTAG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc3	CAAGCAGAAGACGGCATACGAGAT**AGGCAGAA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc4	CAAGCAGAAGACGGCATACGAGAT**TCCTGAGC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc5	CAAGCAGAAGACGGCATACGAGAT**GGACTCCT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc6	CAAGCAGAAGACGGCATACGAGAT**TAGGCATG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

p7_bc7	CAAGCAGAAGACGGCATACGAGAT**CTCTCTAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc8	CAAGCAGAAGACGGCATACGAGAT**CAGAGAGG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc9	CAAGCAGAAGACGGCATACGAGAT**GCTACGCT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc10	CAAGCAGAAGACGGCATACGAGAT**CGAGGCTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc11	CAAGCAGAAGACGGCATACGAGAT**AAGAGGCA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc12	CAAGCAGAAGACGGCATACGAGAT**GTAGAGGA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc13	CAAGCAGAAGACGGCATACGAGAT**GCTCATGA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc14	CAAGCAGAAGACGGCATACGAGAT**ATCTCAGG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc15	CAAGCAGAAGACGGCATACGAGAT**ACTCGCTA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc16	CAAGCAGAAGACGGCATACGAGAT**GGAGCTAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc17	CAAGCAGAAGACGGCATACGAGAT**GCGTAGTA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc18	CAAGCAGAAGACGGCATACGAGAT**CGGAGCCT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc19	CAAGCAGAAGACGGCATACGAGAT**TACGCTGC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc20	CAAGCAGAAGACGGCATACGAGAT**ATGCGCAG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc21	CAAGCAGAAGACGGCATACGAGAT**TAGCGCTC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc22	CAAGCAGAAGACGGCATACGAGAT**ACTGAGCG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc23	CAAGCAGAAGACGGCATACGAGAT**CCTAAGAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
p7_bc24	CAAGCAGAAGACGGCATACGAGAT**CGATCAGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

Supplementary Table 5. Statistical comparisons and associated p-values

p-values associated with figure 2 panel C

Comparisons	p-values
Proteobacteria and Firmicutes	3.2205-06
Proteobacteria and Cyanobacteria	0.0086
Firmicutes and Actinobacteria	4.3440e-09
Firmicutes and Cyanobacteria	0.0094
Firmicutes and Environmental samples	5.0536e-06
Actinobacteria and Cyanobacteria	7.5079e-06
Cyanobacteria and Environmental samples	0.0141

p-values associated with supplemental figure 2 panel D

Comparisons	p-values
Proteobacteria and Firmicutes	0.0021
Firmicutes and Actinobacteria	0.00071
Firmicutes and Environmental samples	0.0052

p-values associated with figure 2 panel D

Comparisons	p-values
Other and biosynthesis	0.0068
Biosynthesis and regulatory	0.0015

Supplementary Table 6. Media, Buffers and cell free mastermix compositions

Media	Composition
2xYT	Tryptone 16g/L, Yeast extract 10g/L, NaCl 5g/L
R5A	Sucrose 100g/L, MOPs 21g/L, $MgCl_2 6H_2O$ 10.12g/L, Glucose 10g/L, Yeast extract 5g/L, K_2SO_4 0.25g/L, casamino acids, 0.1g/L, add 2mL of trace element solution per liter of media, adjust pH to 6.85 with KOH.
Trace Element Solution	$ZnCl_2$ 40mg/L, $FeCl_3$ $6H_2O$ 200mg/L, $CuCl_2$ $2H_2O$ 10mg/L, $MnCl_2$ $4H_2O$ 10mg/L, $Na_2B_4O_7$ $10H_2O$ 10mg/L, $(NH_4)_6Mo_7O_{24}$ $4H_2O$ 10mg/L

Buffers	Composition
S30A	14 mM Mg-glutamate, 60 mM K-glutamate, 50 mM Tris base (pH 7.7 - titrate with acetic acid) + supplement 1 M DTT to make 2 mM DTT concentration before use
S30B	14 mM Mg-glutamate, 60 mM K-glutamate, 5 mM Tris base (pH 8.2 - titrate with 2 M Tris base solution) + supplement 1 M DTT sol to make 1 mM DTT concentration before use

Cell free components	Composition
Amino acid mix	RTS Amino Acid Sampler kit (6 mM of each amino acid, except leucine which is 5 mM)
Energy Buffer	700 mM HEPES (pH 8), 21 mM ATP, 21 mM GTP, 12.6 mM CTP, 12.6 mM UTP, 2.8 mg/mL tRNA, 3.64 mM CoA, 4.62 mM NAD, 10.5 mM cAMP, 0.95 mM Folinic acid, 14 mM Spermidine + 420 mM 3-PGA

Cell Free reaction	Stock concentration	Final concentration	Volume
S. albidoflavus lysate	NA	NA	9.5uL
Mg-Glutamate	500mM	9mM	0.51uL
K-Glutamate	5M	150mM	0.85uL
Amino Acid mix	6mM each*	1.5mM each	7.13uL
Energy Buffer	14x	1x	2.04uL
PEG-8000	40%	2%	1.43uL
Water	NA	NA	0.08uL

*Except for Leucine which is 5mM

This reaction setup yields enough assembled mastermix to carry out 2 reactions. Use 7.5uL of assembled reaction mix to assemble 10uL reactions by adding plasmid library for multiplexed transcriptional measurements or broccoli plasmid and DHFBI-1T for broccoli expression.

References

(1) Hong, H. J., Hutchings, M. I., Hill, L. M., and Buttner, M. J. (2005) The role of the novel fem protein VanK in vancomycin resistance in Streptomyces coelicolor. *Journal of Biological Chemistry 280*, 13055–13061.

(2) Huang, J., Shi, J., Molle, V., Sohlberg, B., Weaver, D., Bibb, M. J., Karoonuthaisiri, N., Lih, C.-J., Kao, C. M., Buttner, M. J., and Cohen, S. N. (2005) Cross-regulation among disparate antibiotic biosynthetic pathways of Streptomyces coelicolor. *Molecular Microbiology 58*, 1276–1287.