## **Supplementary Information**

# Isolation of a Non-genomic Origin Fluoroquinolone Responsive Regulatory Element Using a Combinatorial Bioengineering Approach.

Santosh Kumar Srivastava<sup>1</sup>, V. Rajesh Iyer<sup>1</sup>, Tamoghna Ghosh<sup>1</sup>, Paramesh Ramulu Lambadi<sup>1</sup>, and Ranjana Pathania<sup>1</sup> Naveen Kumar Navani<sup>1,\*</sup>,

<sup>1</sup> Department of Biotechnology, Indian Institute of Technology Roorkee, Uttarakhand, India, 247 667

## **Supplementary Table 1.**

Bacterial strains used in the study.

Genotype and description of each strain as acquired from EcoCyc (http://ecocyc.org/).

Strains	Genotype and description
E.coli DH5a	F <sup>-</sup> , deoR, endA1, gyrA96, hsdR17 (rK <sup>-</sup> /mK <sup>+</sup> ), recA1, phoA, relA1, thi-1,
	$\Delta(lac ZYA-argF), U169\phi80dlacZ\DeltaM15\lambda^{-}, supE44$
E. coli MG1655	F <sup>-</sup> λ <sup>-</sup> ilvG- rfb-50 rph-1
E. coli RFM443	(rpsL galK2 Δlac74) streptomycin resistant
E. coli BW25113 (Wild	<i>F</i> -, $\Delta$ (araD-araB)567, $\Delta$ lacZ4787(::rrnB-3), $\lambda$ <sup>-</sup> , rph-1, $\Delta$ (rhaD-rhaB)568,
type)	hsdR514
E. coli BW25113 Δcra	<i>F</i> -, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), $\lambda^{-}$ , rph-1, Δ(rhaD-rhaB)568,
	hsdR514, $\Delta cra$ (deletion mutant of Catabolite repressor activator)
E. coli BW25113 $\triangle cueR$	<i>F</i> -, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), $\lambda^{-}$ , rph-1, Δ(rhaD-rhaB)568,
	hsdR514, $\Delta cueR$ (deletion mutant of Cu efflux regulator)
E. coli BW25113 ∆dksA	<i>F</i> -, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), $\lambda^{-}$ , rph-1, Δ(rhaD-rhaB)568,
	hsdR514, $\Delta dksA$ (deletion mutant of
E. coli BW25113 ΔyaiA	<i>F</i> -, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), $\lambda^{-}$ , rph-1, Δ(rhaD-rhaB)568,
	hsdR514, $\Delta yaiA$ (deletion mutant of OxyR-dependent induction of
	expression by hydrogen peroxide)
E. coli BW25113 ΔykgD	<i>F</i> -, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), $\lambda^{-}$ , rph-1, Δ(rhaD-rhaB)568,
	hsdR514, $\Delta ykgD$ (deletion mutant of redox-sensitive transcriptional
	activator)
E. coli BW25113 ΔnhaR	<i>F</i> -, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), $\lambda^{-}$ , rph-1, Δ(rhaD-rhaB)568,
	<i>hsdR514</i> , $\Delta nhaR$ (deletion mutant of "Na <sup>+</sup> /H <sup>+</sup> antiporter Regulator)
E. coli BW25113 ΔatoC	<i>F</i> -, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), $\lambda^{-}$ , rph-1, Δ(rhaD-rhaB)568,
	hsdR514, $\Delta atoC$ (deletion mutant of antizyme protein inhibitor of

ornithine decarboxylase)
F-, $\Delta(araD-araB)$ 567, $\Delta lacZ4787(::rrnB-3)$ , $\lambda^{-}$ , rph-1, $\Delta(rhaD-rhaB)$ 568,
hsdR514, $\Delta fadR$ (deletion mutant of Fatty acid degradation Regulon)
F-, $\Delta(araD-araB)567$ , $\Delta lacZ4787(::rrnB-3)$ , $\lambda^{-}$ , $rph-1$ , $\Delta(rhaD-rhaB)568$ , hsdB514 Ager (deletion mutant of acr gene involved with Agridine
transport)
transport) $E = 4(araD - araP)567 = 4(araZ4797(mmP - 2))^{2} - mnh = 1 = 4(arbaD - arbaP)568$
$hsdR514$ , $\Delta mprA$ (deletion mutant of <b>m</b> ultidrug <b>r</b> esistance <b>r</b> egulator)
F-, $\Delta(araD-araB)567$ , $\Delta lacZ4787(::rrnB-3)$ , $\lambda^{-}$ , rph-1, $\Delta(rhaD-rhaB)568$ , hsdR514, $\Delta fur$ (deletion mutant of Ferric Uptake Regulation)
$F_{-} \Lambda(araD-araB)567 \Lambda[acZ4787(\cdots rnB-3)]\lambda^{-} rnh-1 \Lambda(rhaD-rhaB)568$
$h_{sdR514}$ ArecA (deletion mutant of Recombination protein A)
$F_{-} \Lambda(araD_{-}araB)567 \Lambda[ac74787(\cdots rnB_{-}3)]^{-} rnb_{-}1 \Lambda(rhaD_{-}rhaB)568$
$h_sdR514$ AcusR (deletion mutant of <b>Cu-sensing regulator</b> )
$F_{-} \Lambda(araD-araB)567 \Lambda[acZ4787(\cdots rnB-3)]\lambda^{-} rnh-1 \Lambda(rhaD-rhaB)568$
hsdR514. Alra (deletion mutant of Leucine-responsive regulatory protein)
F $\Lambda(araD-araB)567$ . $\Lambda(araZ4787(::rrnB-3), \lambda^2, rrh-1, \Lambda(rhaD-rhaB)568$ .
hsdR514 Afis (deletion mutant of <b>f</b> actor for <b>i</b> nversion <b>s</b> timulation)
$F_{-}$ $\Lambda(araD-araB)567$ . $\Lambda(araZ4787(::rrnB-3), \lambda^{-}, rph-1, \Lambda(rhaD-rhaB)568$ .
hsdR514. AarcA (deletion mutant of dual transcriptional regulator
for Anoxic redox control)
F $\Lambda(araD-araB)567$ . $\Lambda(araZ4787(::rrnB-3), \lambda^2, rph-1, \Lambda(rhaD-rhaB)568$ .
<i>hsdR514.</i> $\Delta fnr$ (deletion mutant of primary transcriptional regulator, FNR
activates genes involved in anaerobic metabolism and represses genes
involved in aerobic metabolism)
F-, $\Delta(araD-araB)567$ , $\Delta lacZ4787(::rrnB-3)$ , $\lambda^{-}$ , rph-1, $\Delta(rhaD-rhaB)568$ ,
hsdR514, $\Delta hu$ (deletion mutant of The HU protein is a small DNA-
binding protein that is considered a global regulatory protein and shares
properties with histones, which play an important role in nucleoid
organization and regulation).
F-, $\Delta(araD-araB)567$ , $\Delta lacZ4787(::rrnB-3)$ , $\lambda^{-}$ , rph-1, $\Delta(rhaD-rhaB)568$ ,
hsdR514, $\Delta$ sodA (deletion mutant of superoxide dismutases)
F-, $\Delta(araD-araB)567$ , $\Delta lacZ4787(::rrnB-3)$ , $\lambda^{-}$ , rph-1, $\Delta(rhaD-rhaB)568$ ,
hsdR514, $\Delta$ stpA (deletion mutant of StpA protein Suppressor
of <i>td</i> phenotype A," is a nucleoid-associated multifunctional protein that
acts as a transcriptional repressor, in chromosomal DNA packaging, and
as a chaperone
F-, $\Delta(araD-araB)567$ , $\Delta lacZ4787(::rrnB-3)$ , $\lambda^{-}$ , rph-1, $\Delta(rhaD-rhaB)568$ ,
hsdR514, $\Delta oxyR$ (deletion mutant of oxidative stress regulator)
F-, $\Delta(araD-araB)567$ , $\Delta lacZ4787(::rrnB-3)$ , $\lambda^{-}$ , rph-1, $\Delta(rhaD-rhaB)568$ ,

	<i>hsdR514,</i> $\Delta hns$ (deletion mutant of <b>H</b> istone-like <b>n</b> ucleoid structuring protein)
E. coli BW25113 ∆crp	<i>F</i> -, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), $\lambda$ , rph-1, Δ(rhaD-rhaB)568,
	<i>hsdR514</i> , $\Delta crp$ (deletion mutant of cAMP receptor protein)
E. coli BW25113 ∆marA	F-, $\Delta(araD-araB)567$ , $\Delta lacZ4787(::rrnB-3)$ , $\lambda^{-}$ , $rph-1$ , $\Delta(rhaD-rhaB)568$ ,
	hsdR514, $\Delta$ marA (deletion mutant of multiple antibiotic resistance
	regulatory protein)
E. coli BW25113 ∆baeR	<i>F</i> -, $\Delta$ (araD-araB)567, $\Delta$ lacZ4787(::rrnB-3), $\lambda$ <sup>-</sup> , rph-1, $\Delta$ (rhaD-rhaB)568,
	<i>hsdR514</i> , $\Delta baeR$ (deletion mutant of <b>b</b> acterial <b>a</b> daptive response)
<i>E. coli</i> BW25113 $\triangle cpxR$	F-, $\Delta(araD-araB)567$ , $\Delta lacZ4787(::rrnB-3)$ , $\lambda^{-}$ , $rph-1$ , $\Delta(rhaD-rhaB)568$ ,
	hsdR514, $\Delta cpxR$ (deletion mutant of conjugative plasmid gene expression
	regulatory protein)

## Supplementary Table 2.

Plasmids used in the study

Plasmid Name	Description/Resistance	References
pNYL-MCS11	Plasmid derived from pZE21	2, 3
pNYL-GFP	GFP gene was amplified from pPROBE-TT'-GFP and was cloned in pNYL-MCS11 plasmid vector on BamHI restriction site.	This study
pNYL-ibsC	Genetic locus coding for toxic peptide ibsC was amplified from <i>E. coli</i> genome and cloned in pNYL-MCS11 vector at BamHI restriction site	This study
pNal::ibsC	Reporter plasmid containing FQ responsive regulating element fused with gene encoding toxic peptide IbsC. Vector backbone codes for kanamycin resistance.	This study
pNal::GFP	Reporter plasmid containing FQ responsive promoter fused with GFP.	This study
pRx::ibsC	Control plasmid containing pNYL-MCS backbone without FQ responsive promoter sequences.	This study
<i>pNal::ibsC</i> -amp	From vector backbone of <i>pNal::ibsC</i> , kanamycin resistant selection marker gene was replaced by ampicillin resistant gene	This study
<i>pRx::ibsC</i> -amp	From vector backbone of <i>pRx::ibsC</i> , kanamycin resistant selection marker gene was replaced by ampicillin resistant gene.	This study

### **Supplementary Figure S1.**

Alignment of promoter regions of representative prokaryotic stress responsive genes using ClustalW2 online bioinformatics program.

# Supplementary Figure S2. Schematic representation of cloning and screening of combinatorial oligonucleotide library.

A) Commercially synthesized single stranded combinatorial oligonucleotide library were converted in to double stranded DNA fragment and double digested with *XhoI* and *Hind*III. Double digested library were ligated with *XhoI-Hind*III digested fragment of pNYL-ibsC/GFP plasmid vector and transformed in *E. coli* RFM443 (Details are available in material and method section). B) To screen the inducible regulatory elements, a live and dead screening strategy has been adopted. Bacterial colonies were selected from the transformed plate and inoculated individually into 96 well plate containing 200 µl LB broth. The grown cultures were then replica plated on a 96 well plate which contained various toxicants including DNA damaging [fluoroquinolones, mitomycin C], protein damaging [ethanol] and heavy metals [sodium arsenite, cadmium chloride, lead chloride] at sub-lethal concentrations and absorbance after incubation was measured at 600 nm. C) The presence of toxicants induces the expression of the toxic peptide IbsC which leads to the killing of the bacterial cells. The IbsC toxic peptide based reporter system selectively screens out the clones harboring functional combinatorial promoters against leaky and constitutive promoters.

## **Supplementary Figure S3.**

*E. coli* BL21 DE3 strain harboring pET28a-hns was induced with 1mM IPTG Analysis of the purified H-NS protein on 12.5% sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE).

## **Supplementary Figure S4.**

Nucleotide sequence of clones representing constitutive promoters (a,b,c,d,e,f,g and h) covering a wide range of GFP fluorescence intensity.

### **Supplementary Figure S5.**

Multiple sequence alignment of FQ-responsive regulatory element with the selected stress promoters from *E. coli* genome. FQ-responsive regulatory element displayed no homology with stress promoter sequences encoded in *E. coli* genome.

## **Supplementary Figure S6.**

Response of FQ responsive promoter in bacterial host lacking efflux pump *tolC* gene and wild type *E. coli* strain in the presence of nalidixic acid. A-B) Wild type  $\Delta tolC$  bacterial strains harboring *pNal::ibsC-amp* (test) and *pRx::ibsC-amp* (control) plasmid was induced with nalidixic acid.

## **Supplementary Figure S7.**

Growth pattern of various stress regulator gene knock out strains of *E. coli* harboring p*Nal::ibsC-amp* and p*Rx::ibsC-amp* fusion system. The growth assays were carried out in duplicate for the *E. coli* mutant strain harboring p*Nal::ibsC-amp* plasmid (test) and p*Rx::ibsC-amp* plasmid (control).

**Supplementary Figure S8.** *In-silico* DNA curvature analysis of FQ responsive (Nal) promoter and their mutant variants at spacer region ('all A' and 'all T').

	10	20	30	40	50	60	70	80
gi/1-60		CGA	TGACTTTAGGCG.	ATCA	ATATAAGATC	G CCGGGCCA	GCAAAGAACT	GCACCCTCCGG
gi.LpxAd/1-60		CGA	T <mark>GACTTTAG</mark> GCG.	ATCA	ATATAAG <mark>A</mark> TC	G CC <mark>GGGCC</mark> A(	GCAAAGAACT	GCACCCTCCGG
gi.EcfF/1-60		A	AGCATACGGGCG.	A <mark>T</mark> GA	CAAATGCAAA	ACTGC <mark>CTGAT</mark> G(	GCTACGCTTA	TCAGGCCTGGA
gi. Yiis/1-60		GCA	AATGACAATACC	C <mark>G</mark> GA	AAATCCTTCT	GAACT CTT CAC	CTTAAGCAATA	TCAAAAAAA
gi.Rsebp3/1-60		A	AGTGGCGGCGCGCG	T <mark>T</mark> GA	AAGAAATTTC	GAACCCTGAGA4	ACTTAATGTTG	TCAGAAGAACT
gi.FusA/1-60		GATAAATCCA	T <mark>GGCTCTGC</mark> GCC	T <mark>G</mark> GC	GAACGAA <mark>C</mark> TT	[ <mark>C-TGATG</mark> -(	CTGCA GAAA	ACAAAGGT
gi.Ybap/1-60		CGCGAC <mark>C</mark>	CGTGGGCGGCAC	A <mark>G</mark> GT	GAGCCAA <mark>C</mark> TT	<mark>C-GCTAC</mark> -(	CAAAACTGGTC	GAACAGGTG
gi.SbmA/1-60		TCGAAACA	ATTCTTATGGTC.	A <mark>G</mark> GC	TGGCGAA <mark>C</mark> TA/	4 <mark>G - CGCCT</mark> - 1	GCTATGGGTC	ACAATGGG
gi.CutC/1-60	C CATT	CTTGAGGATG	TGCGCACGGTCC	G <mark>C</mark> GA	ATTAGGTTTT	C C - TGGACTO	GTGACGGGCG	
gi.Psdp/1-60	GTG.	AGTTC	<b>GGCCTCTGG</b> CAC	C <mark>T</mark> GG	AGCCGGAACA	A A - TCACT	CAGGGCTT	TGTCGAATTCC
gi.EcfA/1-60	· · · · AA · · · TACA	GACAAAGAG <mark>C</mark>	AT CT G C G A A A A A	T <mark>T</mark> G	C/	A <mark>C - GCGGG</mark> A1	GTTCTGGCTG	ATGCTGCTT
gi.Dsbc/1-60		- ACCGAGCGT	<b>CTGCGGCA</b> ACTT	CATC	AACAGCATCA	<mark>C C - CGCGG</mark> G(	CGTGATGTCTG	AA AAGA
gi.RpoDp4/1-60		TACAAATA	ATGCTGCCACCC	T <mark>T</mark> GA	AAAACTG <mark>T</mark> CG/	A T - <mark>GTGGG</mark> A(	CGATATAGCAG	AT AAGAA
gi.GrpE/1-60	GATT	GATGACAATG	TGAGTGCTTCCC	T <mark>T</mark> GA	AACCCTGAAA	C T - GATCCCC	CATAATAAGCG	A
gi.FkpA/1-61	GAGACAGCACT	TCATTTCGC	GT CAT CGAAACT	AATT	TAAAC <mark>A</mark> AA/	A AGA <mark>GTCT</mark> G/	<b>AA</b>	ATAG
gi.Ecfl/1-60	CCAGAACACCCAT	CAGGCGCGTC	CT CAT CGGC TAC	G <mark>a</mark> tg	TAAAA/	A <mark>T</mark> G <mark>GGTCT</mark> GO	GAA	ATGA
gi.BaCap/1-60	AAAGAAAATAATT	AATTTTACAG	<b>CTGTTAAAC</b> CAA	A <mark>C</mark> GG	TTA <mark>T</mark> AA <mark>(</mark>	C CTGGTCAT		ACG CAG
gi.EcfH/1-60	AACGGAGCGGCCC	AGCTTACCTG	CCATTGCACTAA	A <mark>T</mark> AC	TGA <mark>T</mark> AA1	I <mark>GTTGTCT</mark> T -		AACGGC
gi.EcfK/1-60	CT	GCTGTTCCTT	GCG ATCGAA	AAGA	T CAAGGG <mark>C</mark> GG/	A <mark>C</mark> C <mark>GGTAT</mark> C	CGAGCGGGTTC	AAGAC
gi.EcfG/1-60	TTATGT	CACATAAGTG	ACGATGAACGGC	GAAC	TTAATGC <mark>G</mark> AT <mark>(</mark>	<b>T</b> TTTTT	GT	CAGTAGATA
gi.HtrM/1-60		GATT <mark>C</mark>	GGATGTGATGGT.	A <mark>T</mark> GA	TTACAGA <mark>C</mark> AT	T <mark>C</mark> GTGTCTG#	GATTGTCTCT	GACTCCATA
gi.Lpxp/1-60	T	GTCCTGACTA	GTCTTTACAC	. TCT	TTACAGG <mark>A</mark> AC	C AT <mark>tgtcg</mark> t/	ACATGATGGCC	CAACCA
gi.EcfM/1-60		GACCTTTGTT	ACAATTAGATTC	AATT	TGAATTT <mark>A</mark> TG	T TTTTGAATO	GCTTTCTTATC	TCAC
gi. Yfey/1-60		GCT - TGTT	TTTATGAAGTAA	AAGA	ATAACGG <mark>C</mark> AC <mark>I</mark>	「.. <mark>TTTTGGT</mark> G#	ATTTGCACTC	CA AGCAA
gi.EcfD/1-60		T.TTTA	TT <mark>CGCGAAC</mark> GTG	AATA	AT C C G G G A A C /	A TTTCGGCC/	AAGCCTGATC	TAAGCGTTG
gi.RpoE2/1-60		<mark>T</mark>	<mark>Gacaaac</mark> aaaaa	CAGATGC	GTTACGG <mark>A</mark> AC	T T T ACAAAAA	ACGAGACACTC	TAACCCTTTG-
gi.RseA/RseB/1-60		<mark>T</mark>	GACAAACAAAAA	C <mark>A</mark> GATGC	GTTACGG <mark>A</mark> AC	T TTACAAAAA	ACGAGACACTC	TAACCCTTTG-
gi.RpoH3/1-59		- TGCGTAATT	TATT CACAAGCT	Т <mark>G</mark>	CTT G <mark>A</mark> AC	T TGTGGATAA	AATCACGGTC	TGATAAAAC
gi. YeaY/1-60		· · · TTTATT	ACGTAACAACGT	C <mark>G</mark> CATGG	AAG A <mark>A</mark> AC	T T C <mark>CGGGC</mark> A#	AGAATGAATC	TT AAGAG
gi.HtrA/1-57		· · · TTTACCT	TT <mark>GCAG</mark> AAACTT	T <mark>A</mark>	TTCG <mark>G</mark> AA <mark>(</mark>	C TCAGGCTAT	AAAACGAATC	TGAAGAACA
gi.SurA/1-58	GGCCAC	GCCAATGAT	TTTCCCCAA - TA	C <mark>T</mark> G	CATATTC <mark>C</mark> (	C <mark>Caaatcg</mark> a(	CACACGGA	· · · · · · · · TAT · · ·
gi.SixA/1-60	CGT	GGGGAAAGTT	TTTGGAAAA - CA	A <mark>C</mark> T	GCAACTG <mark>A</mark> <mark>(</mark>	C CTGCAATAA	AGAAGGTCAA-	AGCTAT
gi.RecA/1-60	••••CT••••TGT	GGCAACAATT	TCTACAAAA - CA	C <mark>T</mark> T	GATACTG <mark>T</mark> /	A TGAGCATA(	CAG-TATAAT-	TGCTT
gi.EcfE/1-60		CAGGAATT	AAGGACAGC	G	GTCATTTAAT	CCAGGACAG	CGGTGGTATTT	TAGATCGTATTG -
gi.EcfJ/1-60		GCGGGAAATT	CGATAAATAGCA	CAT	ATGATTAAAA	C TCAGACCC/	AGTGGTCGGA	TCACC





Lane-1 Broad range protein marker Lane-2 Purified fraction1 Lane-3 Purified fraction2 Lane-4 Uninduced Lane-5 *pET28a*-hns cellular extract induced with 0.5 mM IPTG

#### Clone-a

CTTTTGTTTACGTCTTCC<u>CTCGAGATCAGTCAAATACAGCGCAACTCGTAAAAAGGGGGAGTTTCCAACAATCCAGGCAGTCATAT</u> <u>CTAAGCTTAGATATGACGCACCGTTTATTCCGAGCCTGAGGGTTTTTCGGGGGTTTGTCCTCTAATGACTTGATCTCGAGATCAAGT</u> <u>CATAAAAGTTGAAACCATCAAAAAGGCGCAGCCTCAGATTATCACCGGGGGTCATATCTAAGCTTAGATATGACACCCTGTTGGT</u> <u>GGGACCTACTCCCTTTTTTAGAGTTGACGTCAAATTGACTTGATCTCGAGATCAAGTCATTATAGTGGCAACTATGAAAAATACGA</u> <u>CTGGTCGCACCAACCGGCGAGTCATATCTAAGCTT</u>GATATCGAATTCCTGCAGCCCGGGGGATCCATAAGGAGGAAAAACATATG AGTAAAGGAGAAGAACTTTTCACTGGAGT

#### Clone -b

TGGGTTTTTTTATCTTTC<u>CCTCGAGATCAGTCAATTCCCGTTAAACCACGAAAAATAACAGGTTTCGTATAAAACGCGCGGGTCATA</u> TCTAAGCTTAGATATGACGCCCGGGTAAATCTGAAAGTGCAATTTTTTACTAGTTCGCCAGGTATTGACTTGATCTCGAGATCAAG TCATATTTAAACAAACCCGGAAAAAGAGCAATGGTCAAACTATCCGGGGCGTCATATCTAAGCTTAGATATGACCCCCAGGTTAG TCAGACACCCCCCCTTTTTTGTGGTTGAAACCGTTTTGACTTGATCTCGAGATCAAGTCAAGTCAAGTCAAAAAAACCCG GCATATCGAATTTTAAGGTAAGTCATATCTAAGCTTGATATCGAATTCCTGCAGCCCGGGGGATCCATAAGGAGGAAAAACATAT GAGTAAAGGAGAAGAACTTTTCACTGGA

#### Clone- c

#### Clone-d

 $GTCCCCGGGGGTTTTTATTCTTC\underline{TCGAGATCAGTCATTTTCTTATAAACTCATAAAAATTGAACTATTCTTAACTAAAGATAGGTCAT\\ \underline{ATCTAAGCTTAGATATGACTCCCCGTTAAATCCGACCCGGGGGCATTTTTGATAGTTTACGGGAAAATGACTTGATCTCGAGATCAA\\ \underline{GTCAATTTAAGACAAACCCTTAAAAAGCGTATGGCTCGCACAAACCGGGCGGTCATATCTAAGCTTGATATCGAATTCCTGCAGC\\ CCGGGGGATCCATAAGGAGGAAAAACATATGAGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTG\\ \\$ 

#### Clone- e

TAGGGTTTTTGTCTTA<u>CTCGAGATCAGTCAAATCGCGCGACCAAAAAAAAGGGGTGCCGGTCCAATTTACCTCCCCGTCATATCTA</u> AGCTTAGATATGACCCCTAGGAATGTCAGACCAGGGCTCTTTTTATTTGTTGAATATGTTATGACTTGATCTCGAGATCAAGTCAA TTGCGGCTCAACCAGGAAAAAGCGAGCAAGTCGGATCTTCACGAGTGTCATATCTAAGCTT</u>GATATCGAATTCCTGCAGCCCGGG GGATCCATAAGGAGGAAAAACATATGAGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGA TGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGT

#### Clone f

#### Clone g

#### Clone-h

 $CGGGGGTTTTTTCGGTCTTAA \underline{CTCGAGATCAGGTCATTATCGTTGGAACTATTAAAAATTTTTGTGTTCGTAAATACCCATTCGTCATAGGTCAGACGTCTTAAGGTTAGACTCGGAGAAAATCCGACGTCAGACGTTTTTCGTTGTTTGCCAACATATGACTTGATCTCGAGATCAAGGTCAGGGGTACCCTTTATGATTTTCATACTCGCTCTATAAATCAGATAGGATCATATCTAAGCTTGATATCGAATTCCTGCAGCCCGGGGGATCCATAAGGAGGAAAAAACATATGAGTAAAGGAGAAGAACTTTTCACTGGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAAATTTTCGTCGTGGAGGAGAGGGT$ 

	1 10	20	30	40	50	60	70	80	90	100	110	) 114
gi.NAL	GTCATGGTATAAA	TACTGATTTTT	TTTAGTG	TTTGTATATCCT	CATGTCAT	+	+	+	+	+	+	
gi.accA	GGATATTGACGAAG	CCCGGGAAAAA	TATGCTC	GCGGGCTTGCTAT	CTCGCTGA	CGGACAGGCAA	ATTGATGACC	AGCTTT-T	AAACCGACTC	CGTCAGTCTC	;	_
gi.mtr gi.ondP	IHIHIHCHG	UUUUU-HIIII TCCTTOOTCOO	TECECCO	і-бббСТТТТТТТТ Статаотостат	CIGICITI	IGIHCICGIGI FCOOOCCCTOO	HCIGGIHCHG	16CHH16C	НІННСННССС	HGICGCHCIF	ПТТТТСТС	
gi,guaBA	TGGGCGGTG	TCGTCTTTGAG	TGTI	AAGTACCAGTGA	CCGGAAGC	FGGTTGCGTGA	AATTAGAAAT	TTCGCCGCTG	ATCCARACCT	GTCCCATCTO	ATGCTC	
gi.pdx	GTTGTATCATT	ACGTATCCTTA	TACC	GAAATCTTCGCA	AGTATGCC	FGGCCGCGAGA	TTATGGCACA	CTTGTCCG	GTTAACTCTC	GTCTCATACE	IGGTAA	
gi,riDHP1 gi,rho	GTAAGGGAATT	TCATATTCAGA	LGHC· TGCC·	-666766678888	BETGGACG	LETTETEE===	GTCATGCTAA	HTHHGLLH CTTAGTGTTG	ACTTCGTATI	. HGCHIIHICI "AAACATACCI	. GGHGHH (ТАТТА	1
gi.ahpC	TGTAAGGTA	AAACTTATCGA	TTTGATĂ	TGGAAACGCATT	AGCCGAAT	CGGCAAAAATT	GGT-TACCTT	ACATCTCATC	GAAAACAC	GGAGGAAGTA	ITAGAT	
gi.lpd	GTTATTAG	CGAATAGACAA	ATCGGTT	GCCGTTTGTTGT-	TTAAAAAAT	IGTTAACAATT IGTTGICACIC	-TTGTAAAAT	ACCGACGGAT	AG-AACGACC	CGGTGGTGGTGGT	TAGGGT	ſ
gi.phe	ATTGAG	TGTATCGCCAA	CGCGCCT	CGGGCGCGTTTT	TTGTTGACI	AGCGTGAAAAAC	AGT-ACGGGT	ACTGTACTAA	AGTCACTTAA	IGGAAACAAAC	ATGAAF	ì
gi.fabA	CATTACGTT	G-GCTGAACTG	GTTTATT	CGAACTGATCGG	ACTTGTTC	AGCGTACACGT	GTT-AGCTAT	CCTGCGTGCT	TCAATAAAAA	AAGGCTTACE	IGAGA	
gi,dsda	LINHHUII TACGGAC	6161C111C16 TC-6CTTTTCT	GGCTAGT	ILGLHLIHILLHGI GAAGCAATATTGCI	ACTIGHAN	11 HG I CGC GCTCAA-TGAA	GIH-HUUUHI ACATTTAAAT	ACTATACGAC	1 H I LGLH I H I AGCGACATTT	ATC6CTTCA	THHHUI	i G
gi.rbs	AACGTT	TC-GAGGTTGA	TCACATT	CCGTAACGTCAC	GATGGTTT	TCCCAACTCAG	TCAGGATTAA	ACTGTGGGTC	AGCGAAACGT	TTCGCTGATO	GAGAAF	i
gi.htrA	AGACGAAC	AATAAATTTTT	ACCTITI	SCAGAAAACTTTAG CTCOTOTOTOTCTT	TTCGGAAC"	TCAGGCTATA	AAACGAAT	-CTGAAGAAC	ACAGCAATTI	TGCGTTATCI	GTTAAT	ТС
gi.cps	ATATAAT	TCTTAAAAATA	GCCAATTI	ACCGAATTGTTAT	CTT-GCCT	GCTATTCCGTT	AGCT-GTAAC	ACTTCCTCCT	GCATTATTGG	AAAGCCAATA	ITTCAAT	
gi.rpmFP2	ATTATTGG	CAAAAGGCA	ACCACAG			CTATGACGTTA	CAAA-GTTAA	TATGCGCGCC	CTATGCAAAA	GGTAAAATTA	ICCCCTG	iac .
gi.fis	AATTCTTT	GATCCATCTCA	GAGGATT	GTCAAAGTTTGG	CCTTTCAT	CTCGTGCAAAA	AATGCGTAAT	ATACGCCGCC	TTGCAGTCAC	AGTATGGT-0	ATTT:	1
gi.nusA	TTTTAC	GTCCCGTCTCG	GTACACCI	AATCCCAGCAGT	ATTTGCAT	ITTTTACCCAA	AACGAGTAGA	ATTTGCCACG	TTTCAGGCGC	GGGGTGGAGC	AGCCT	
gi.rimL ai ppoSP2	GAGATTTAAT	ATCCCGTCTGA	AAAACGAI GCACAAA	HAATCCTCGCTTTI	DJT6TCC	-TTGTTCGATA CCCAACCGACA	ААСАСААТАА АТТАССТ	АСТТБАТ-СА ТТСТБАБТСТ	TCGGGTGGGC	AGAAAAGGAAU	iGTTCC ICAAAAT	r
gi.aspA	AATTA	AGGGGTAAAAA	CCGACAC	TAAAGTGATCCA	GATTACGG	TAGAAATCCTC	AAGCAGCATA	TGATCTCGGG	TATTCGGTCG	ATGCAGGG-C	ATAATC	G
gi.flaA	AACGTC	AGAGGTAGCAC	CGTAATC	GCGTCTTTTCCC		TGCGCTCAAG	ACGCAGGATA	A-TTAGCCGA	TAAGCAGTAG	CGACACAG-C	AAGACC	GC
gi.umu gi.uvrBP2	GATTGAC	AGCGGAGTTTA	CGCTGTA	ICAGAAATATTAT	GGTGATGAI	ACTGTTTTTTT	ATCCAGTATA	ATTTGTT-66	CATAATTAAG	TACGACGA-0	1CHGH11 111000	Г
gi.crp	AACGGA	AGGCGACCTGG	GTCATGC	rgaagcgagacac	CAGGAGACI	ACAAAGCGAAA	GCTATGC-TA	AAACAGTCAG	GATGCTACAG	TAATACATTO	iATGTAC	;
gi.fdh⊦ gi ecnDP1	IIIGHHIHHH ACTT	HG   GUUU    ATCAGTAAGTA	GINHINI TTCACCA	CHGGGHHIGHCCCI STTAAATTGATTGI	СНСН І НННІ ВАТСААТА	11G1GGCHIHH FACAGGGAATA	ННGНIGCHIH Атаатттста	CIGINGICGH TTTTATATTA	IGHGCGCGIHI	GCG16H1110 TAATTAACTO	1411H • TATCAC	6
gi.lpp	CGAACGATC	AAAAATAAGTG	CCTTCCC	тсаааааатат	TCTCAACA	гаааааас	TTTGTGTAAT	ACTTGTAACG	CTACATGGA-	-GATTAACTO	AATCTF	iG
gi.fecI	ATAACACTT	AGAAAAAACAAC	ATGTTAAI	AATGTCTATTGG	RAACAATT	111TAT 1111TAT	CAATTGTAAT	GATAACCATT	CTCATATTA-	-ATATGACTA	ICGTGAT	AA CT
gi,lexA	CAGAATTCG	ATAAATCTCTG	GTTTATT	GTGCAGTTTATGG	TTCCAAAA	ICGCCTT	TTGCTGTATA	TACTCACAGO	ATAACTGTA-	-TATACACCO	AGGGGG	icg
gi.araBAD	CCATAGC	ATTTTTATCCA	TAAGATTI	AGCGGATCCTAC-	C-TGACGC	TTTTTATCGCA	ACTETETACT	GTTTCTCCAT	ACCCGTTTT-	TTTGGATGGA	IGTGAAA	10
gi.gut	CAAATTAAA	ATATTTATCTT	TCA-TTT	GCG-ATCAAAA-·	TAACAC	ITTTAAATCTT	TCAATCTGAT	TAGATTAGGT	TGCCGTTTGG	TAATAAAACA	IATAAAT	
gi.ompC	CGATAAA	TGAAACATCTT	AAAAGTT	TAGTATCATATT	CGTGTTGG	ATTATTCTGCA	TTTTTGGGGA	GAATGGAC-T	TGCCGACTGA	TTAATGAGGG	ITTAAT	
gi.hns	AAAATAAAG	AACAATTTTGA	ATTCCTT	ICATTCCTGGCTA	TTGCACA-I	ACTGAATTT	AAGGCTCTAT	TA-TTACCTC	AACAAACCAC	CCCAATATAA	IGTTTGF	í
gi.oppf	TCACCATCA	TGTTATTTCG	CCACATC	ATAATCCTGG-GC	TTGCTGA-I	IGAATAATT	GAAATGATAT	TA-TTAATTC	CA-CTGCCTT	TGGTAGAGGA	AAGTGC	TA
gi.codB	GCACTC	ATTCATATAAA	АААТАТА	TTCCCCACGAAAA	ACGATTGC	ITTTTATCT-T	CAGATGAATA	GA-ATGCGGC	GGATTTTTT	GGTTTCAAAC	CAGCAAF	ie -
gi.emrR	TTATA	TAACCCACAAG	AATCATT	TTCTAAAACAATI	ACATTTAC	TTATTGTCA	CTGTCGTTAC	TA-TATCGGC	TGAAATTAAT	GAGGTCATAC	CCAAAT	G
gi.tolH gi.recA	TATGCATTGC	AGACCTTGTGG	СААСААТ	TCTACAAAACACI	ГСАНСАНСІ ГТСАТ-АС	16TAT6A6	CATACAGTAT	88-TTGCTTC	.880.86880.81	ATTGACTATO	CGGTA7	Γ Γ
gi.pdhR	TAAAAT	GTGCACAGTTT	CATGATT	TCAATCAAAACCT	GTATGGACI	ATAAGGTGAAT	ACTTTGTTAC	TT-TAGCGTC	ACAGACATGA	AATTGGTAAC	iaccaat	r
gi.rha gi.fadl	AAATC	GACGITITIA	CGTGGTT	TCCGTCGAAAAT	TTAAGGTAI	AGAACCTGACC	TCGTGATTAC CCGAC_CTAT	18-111CGCC 861_61666	GTGTTGACGA 9	ICATCAG-GAG GATTTCTAAG	GCCAGT	TH
gi,queA	AAAATTCAT	CGATTATATTC	TATCCAA	AGGGGGGTAAAGG	TTGCAGGG	AGAGCGCC	CCGGCACTAG	ACTACCCGCC	TCTTATTT	TAGTCTGAGT	CAGTGT	ĩč
gi.katE	CTGGCTT	CACTAAACGCA	TATTAAA	ATCAGAAAAA-CI	T-GTAGTT	TAGCCGATT	TAGCCCCTGT	ACGTCCCGCT	TTGCGTGTAT	TTCATAACAC	CGTTTC	CA
gi.aroH	CTCGA	TATCATGGGCC	TTAGTCG	CGAATGTACTAG	AGAACTAG	I GCATTAGCTT	ATTTTTTTGT	TATCATGCTA	ACCACCCG-G	CGAGGTGTGF	ICACACC	Ϊ
gi.relA	CGAGCAGG	TGCCGCAGGTG	TTATGCA	SCAAATTATI	RAAACTGG	ACCTATTCGT	ATAGTTTATG	TATCCTGTAA	CCCTGCAACO	CTGGCTCGGC	ATAGCO	ĥ
gi.tran gi.oxyR	GCCGCTC	CGTTTCTG-TG	AGCAATTI	.GATCAGTCAGA(	ATGCTTGA	IAGGGATA-AT	CGTTCATTGC	TATTCTACCT	ATCGCCATGA	ACTATCGTGO	ICGATG(	ICH GAGG
gi.rpsAP3	CGGCAGCCG	ATGCTTTAGTG	TTGGATT	CACCACCTT	RAGCATTG	AGCAAGTGATT	GAAAAAGCGC	TACAATACGC	GCGCCAGAAA	TTGGCTCTCC	icataa	
gi.uvrCP2 gi.dcoA	НСНННННССН ААБТСТТТБ	IGHGHHHHHH I CG CGCTGGAACCG	HH <b></b> 1HCI CGTTAAA'	CHCCHIIII TCACGCCCTTI	TCTTTTG	СНАНИСТТСТТТТ ПСАТТТСТТТТ	ТНННННСССС ССАСТССТАА	І НН І Н І НССН АСТАВАТСАС	HIGHUIHH TTTTTTTTGT	ICTHTCHGTHU 77786671	ICG11H1	
gi.fadBA	GCCG	AGCGTGATCAG	ATCGGCA	TTCTTTAATC	TTTTGTTT	GCATATTTTA	ACACAAAATA	CACACTTCGA	ICTCATCTGGT	ACGACCAGAT	CACCTT	GCG
gi.groE gi btoGP1	CGATTITG	TGCT-GATCAG	AATTTTT AGCAGGA'	TTCTTTTTCCCCC	CTTGAAGGO	GCGAAGCCTC	ATCCCCCATTT	CTCTGGTCAC	CAGCCGGGAA CTGTTTT <b>-</b> TA	ACCACGTAAC	JOTOC Daan	
gi,ibpB	TGCAATAG	ACCATAAACTG	CAAAAAAA	AGTCCGCTGATA	AGGCTTI	GAAAAGTTCAT	TTCCAGACCC	ATTTTTACAT	CGTAGCCGAT	GAGGACO	ICGCCTE	iat
gi.tau	TTGCTA-AG	AATAAAATCAT	CTGTGCG	TAACGACTAATT		GAATGTTTTA	TTCCTGAATA	CTGCTCCCAT	AACAAGAC	AGGGGGAC	ICAGACA	IAT
gr.gym gi.	ACTCACAAAAAGA	ICAC-GCGTTTA	ATTTGCG	TACGAATTAAAT	TTTCACACI	ACTCTGTAGCA	GATGA-TCTA	ACAATCTGAT	TACAG-AACA	TCGGCAGTAC	.na i nan 288	1000
gi.pabAP1	GGGGATCT	CACCGCACTGC	AGATGAT	ITTTAGCAAAGTG	GTAAGCGA	AGCCGGGGG-AA	TCTGAGTAAA	ATAGCGCGGT	TCTTT-TGTP	CCGGAGCCGC	CATGA:	
gi.uvrH gi.bupB	HIIGIIH GAA	ІСНСННСНСТСС ІАССАБСТСТ-С	GGGINNIO	ICHTTCCHHTHCT TGTTTTGATCTG	GTHTHTTC	11 I CHGG I CHH CGGCAG-CGTG	I I I GI GI CHI GCGGAAGGTT	HH I I HHUUG I GTAAACTGCA	11616-H1C6 CCTC688688	ICCGG I HGCHU ICAAGAGGCCG	:CHIGC 38876C7	атаат
gi,lrp	TGAAT	ATCTGGCAT-G	TTGTACT	AAAAATCGATG	TTTTGCTT	rgacaa-tccc	CTGGTGTTTT	GCGAAAACAT	TCGAGGAAGA	AAAAAAACAC	TATTCT	TATA
gi.argCBHP1	CGGTAAATCT	CGATAAATGGC	GGTAATT	GTTTTTCATTGT	TGACACACI	C-TCTGGTCAT	-GATAGTATC	AATATTCA ATCAGCAA	TGCAGTATTI	ATGAATAAAA TATGAATAAAA	IATACA	
gi.osmBP1	AAC	CATTTCCCTTC	TCTACGG	TGATTTGCAGTT	TGGCAAAT	CATCCGCTCTA	AGATGATTCC	TGGTTGATAA	TTAAGACTAT	TTACCTGTTF	ITTAACF	іст
gi.fepf	GCCGT	CACACCATAAC	CCCATGT	TACTGTGCAATT	TTTCAT'	IGATTGCAGAA	ATATATTGAT	AATATTAT	TGATAACTAT	TTGCATTTGC	AATAGO	GTAA
gi.tes gi.appA	GTCTGAAAAGTT	AACGAACGTAG	GCCTGAT	GCG-GCGCATTAG	CATCGCAT	nn i GCHAAN C-AGGCAA	поттитенит ТСААТААТ	GTCAGATATG	AAAAGCGGAA	ACATATCGAT	GAAAGC	1111 )
gi.hscA	AACAGTTAGACA	ACGAGACGTGG	GAC-GCG	GCG-GCGGATACC	GTGCGTAA	G-CTGCGT	TTTCTCGATA	AACTGCGAAG	CAGTGCCGAA	-CAACTCGAF	IGAAAAF	1
gi,timBP1 gj.groF	HIHHHIHIHTA GGTCCTTG	нньь і нСН і АС ТТСССААБАТТ	-CHIGCAI GATGACAI	1—————————————————————————————————————	GITIGIHTO CCTTGAAAA	ыысннСGT СССТбаааста	1411414467 A-TCCCCATA	нисны і і GCT Атаабсбааб	HCTCCHTTFF - TTAGCGAGE	INGT TCHUTCF ITGAATGCGAF	JGHHGHA 188888	16
gi.otsA	CAGGTGCAAC	TCAGGCATCAT	GGCGACT	G-GCGGGTG-TGC	CGGAT	GTCTGGAGCTG	GCTTGAAATG	ATAACCACCG	CATTACAACA	AAAAAGAGAA	інатаас	:
Consensus	•••••	•••••	•••••	•••••	•••••	• • • • • • • • • • • • • • • • • • • •	•••••	• • • • • • • • • • • •	•••••	•••••	•••••	•••

A









