Supplementary Information

Droplet-based high-throughput single microbe RNA sequencing by smRandom-seq

Ziye Xu^{1,2#}, Yuting Wang^{2,3#}, Kuanwei Sheng^{4,5#*}, Raoul Rosenthal⁶, Nan Liu², Xiaoting Hua⁷, Tianyu Zhang², Jiaye Chen⁸, Mengdi Song², Yuexiao Lv², Shunji Zhang³, Yingjuan Huang², Zhaolun Wang², Ting Cao^{1,4,6}, Yifei Shen¹, Yan Jiang⁷, Yunsong Yu⁷, Yu Chen¹, Guoji Guo², Peng Yin^{4,5*}, David A. Weitz^{4,6*}, Yongcheng Wang^{1,2,3,4*}

¹Department of Laboratory Medicine, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

²Liangzhu Laboratory, Zhejiang University, Hangzhou, China

³College of Biomedical Engineering and Instrument Science, Zhejiang University, Hangzhou, China

⁴Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, United States

⁵Department of Systems Biology, Harvard Medical School, Boston, United States

⁶John A. Paulson School of Engineering and Applied Sciences and Department of Physics, Harvard University, Cambridge, United States

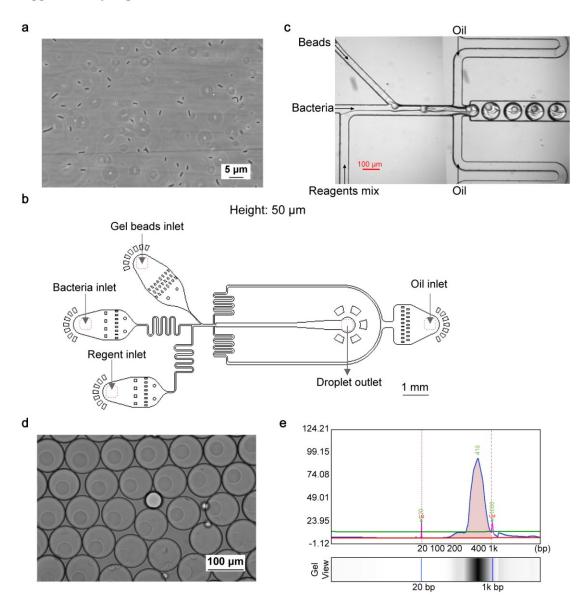
⁷Department of Infectious Diseases, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China

⁸Department of Biomedical Informatics, Harvard Medical School, Boston, United States

#These authors contributed equally to this work

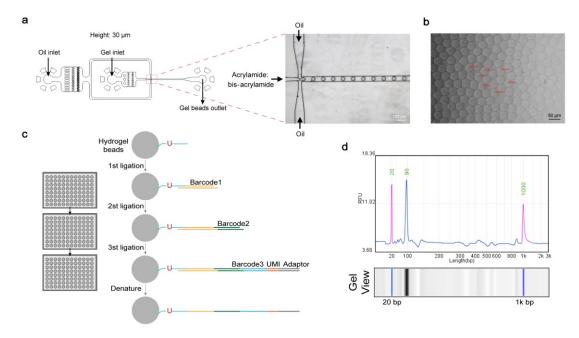
*Corresponding Author: Yongcheng Wang (<u>yongcheng@zju.edu.cn</u>), David A. Weitz (<u>weitz@seas.harvard.edu</u>), Peng Yin (<u>Peng Yin@hms.harvard.edu</u>), and Kuanwei Sheng (Kuanwei.Sheng@wyss.harvard.edu)

Supplementary Figure 1



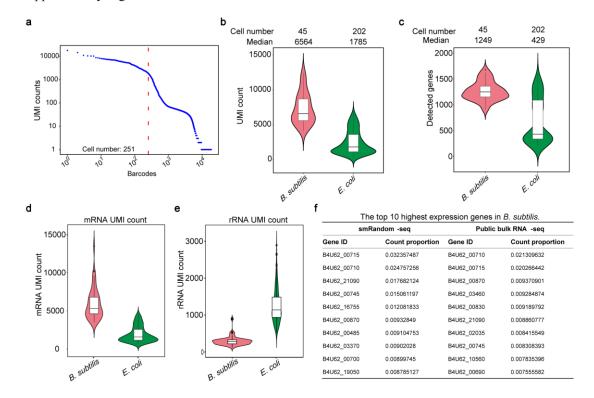
Supplementary Figure 1. The microfluidic platform of smRandom-seq.

a, Image of fixed *E. coli* after reverse transcription and dA tailing. Scale bar: 5 μm. n= 4 independent experiments. **b**, Design of the device for cell, bead and mix reagents encapsulation. Height: 50 μm. **c**, Image of microfluidic barcoding device. Scale bar: 100 μm. **d**, Image of encapsulated droplets. Scale bar: 100 μm. **e**, Electropherogram of amplicon from barcoded single bacterium cDNAs. Lower marker: 20 bp; upper marker: 1k bp.



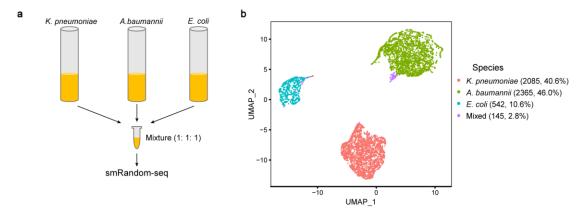
Supplementary Figure 2. Synthesis and quality control of barcoded hydrogel bead.

a, Design of the device for hydrogel beads generation. Height: 30 μm. Hydrogel bead generation and collection. Scale bar: 100 μm. **b**, Image of hydrogel beads. Scale bar: 50 μm. **c**, Barcode beads fabrication procedures. U: uracil residue; UMI, unique molecular identifier. **d**, Electropherogram of released DNA primers after enzymic digestion from barcode beads. Peaks at 20 and 1000 base pairs are gel migration markers.



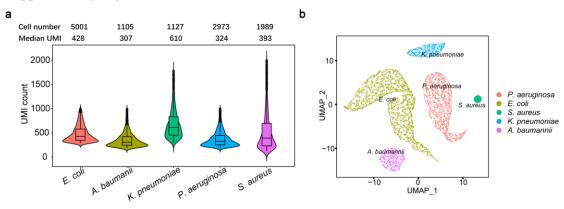
Supplementary Figure 3. Performance of the smRandom-seq on reference bacteria samples.

a, UMI count versus barcode rank plot of the *E. coli* and *B. subtilis* mixture. The knee point (red dashed line) indicates the threshold for *E. coli* and *B. subtilis* mixture (Cell number: 251). b, c, Distribution of UMI counts (b) and detected genes (c) from *B. subtilis* and *E. coli* in the mixture of *B. subtilis* and *E. coli* sample by smRandom-seq without rRNA depletion. *B. subtilis* n= 45 cells, *E. coli* n= 202 cells. d, e, Distribution of mRNA UMI counts (d) and rRNA UMI (e) counts detected from *B. subtilis* and *E. coli* in the mixture of *B. subtilis* and *E. coli* sample by smRandom-seq without rRNA depletion. *B. subtilis* n= 45 cells, *E. coli* n= 202 cells. f, The top 10 highest expression genes sorted by the count proportion in *B. subtilis* datasets of the species mixture by smRandom-seq and the public bulk RNA-seq dataset of the Emp3h_i Exponential growth sample¹. Data in the box plot in **Supplementary Fig. 3b-e** corresponded to the first quartiles (lower hinges), median (center), and third quartiles (upper hinges). Whisker is defined as 1.5* IQR (interquartile range) from the hinge.



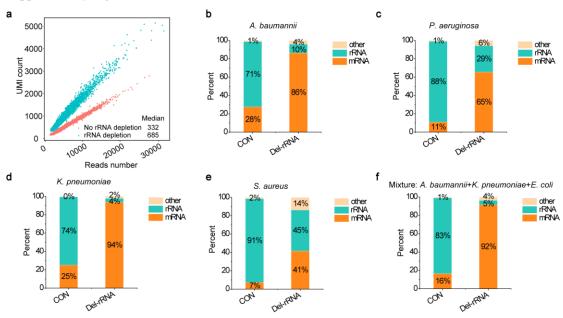
Supplementary Figure 4. Three-species-mixing experiment.

a, Experimental design for three-species mixture including *A.baumannii*, *K. pneumoniae*, and *E. coli* by smRandom-seq. **b**, UMAP plot colored by species, *A.baumannii*, *K. pneumoniae*, *E. coli* and mixed cells, identified *from the* three-species mixture.



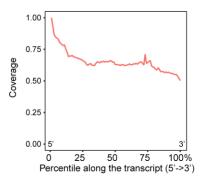
Supplementary Figure 5. Single-species experiments.

a, Distribution of UMI count of different bacterial species datasets by smRandom-seq, including *E. coli*, *A. baumannii*, *K. pneumoniae*, *P. aeruginosa*, and *S. aureus*. The cell numbers of these bacterial species datasets by smRandom-seq were as follows: 5001 *E. coli*, 1105 *A. baumannii*, 1127 *K. pneumoniae*, 2973 *P. aeruginosa*, and 1989 *S. aureus*. Data in the box plot in **Supplementary Fig. 5a** corresponded to the first quartiles (lower hinges), median (center), and third quartiles (upper hinges). Whisker is defined as 1.5* IQR (interquartile range) from the hinge. **b**, UMAP plot colored by species, *E. coli*, *A. baumannii*, *K. pneumoniae*, *P. aeruginos*, and *S. aureus*, identified from single-species experiments. Source data are provided as a Source Data file.



Supplementary Figure 6. Case9-based rRNA depletion experiments.

a, Scatter plot showing UMI count and the number of reads per barcode of *E. coli* samples with (blue) and without (red) rRNA depletion. Each dot corresponds to a putative cell. **b-f**, Representative proportions of transcript categories of single bacterial species, including *A. baumannii* (**b**), *K. pneumoniae* (**c**), *P. aeruginosa* (**d**), and *S. aureus* (**e**), and a mixture of *A. baumannii*, *K. pneumoniae*, and *E. coli* (**f**) with (Del-rRNA) and without rRNA depletion (CON). other, all other RNA classes. Source data are provided as a Source Data file.



Supplementary Figure 7. Average percentile coverage across all transcripts (5'->3'). Source data are provided as a Source Data file.

Supplementary Figure 8 а b 30.5 30.0 6000 GFP fluorescence/OD600 29.5 5000 _ට 29.0 4000 28.5 3000 28.0 2000 27.5 1000 27.0 E1 E2 E3 E4 E1 E3 С d 120 150 = 0.0018 Median UMI count 95 Median UMI count p = 0.03330

Supplementary Figure 8. Quantitativeness of gene expression in smRandom-seq by GFP expression model.

2 3 4 5 qPCR relative expression

5

6000

5000

3000

2000

4000

GFP fluorescence intensity/OD600

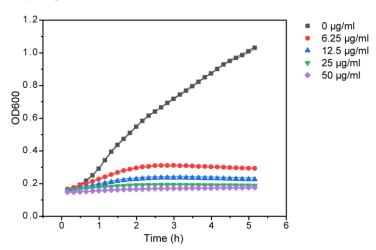
a, b GFP fluorescence/OD600 values (a) (n=6) and Ct values from qPCR (n=3) (b) of E. coli samples with different GFP expression induced by different amount of propionate (0, 6.25, 25, and 100 mM). Data in a, b were expressed as mean ± SEM. c, Comparison of GFP fluorescence values and the median number of UMIs. d, Comparison qPCR relative expression and the median number of UMIs. Relative quantification in qPCR is calculated by comparing the expression levels of GFP in other groups (E2, E3 and E4) against the expression levels of GFP in control (E1). R presents Pearson correlation coefficient. The shaded area represents the 95% confidence interval. p-value is the significance level of the t test and calculated alongside the Pearson coefficient as implemented in the R package ggscatter. Source data are provided as a Source Data file.

Supplementary Figure 9 a Repeat 1 SmRandom-seq Repeat 2 SmRandom-seq Repeat 2 Repeat 1 Repeat 1

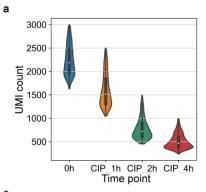
Log10 (Counts+1)

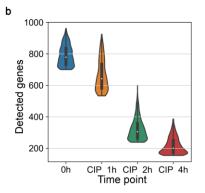
Supplementary Figure 9. Evaluation the technical reproducibility of smRandom-seq.

a, Experimental design for technical reproducibility evaluation. **b**, Scatter plot showing correlation of gene expression (Log10(Counts+1)) for two repeated E. coli samples (Repeat 1 and Repeat 2) applied with smRandom-seq separately. R presents Pearson correlation coefficient. p-value is the significance level of the t test and calculated alongside the Pearson coefficient as implemented in the R package ggscatter. Source data are provided as a Source Data file.



Supplementary Figure 10. Growth curves (OD 600) of *E. coli* upon 0, 6.25, 12.5, 25, and 50 μg/mL CIP. CIP: ciprofloxacin. Source data are provided as a Source Data file.





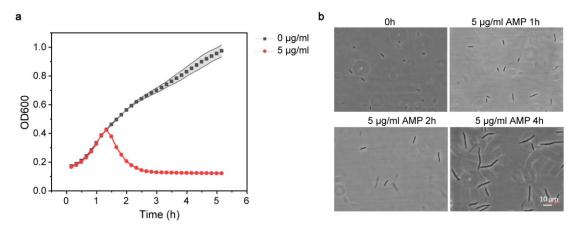
С

The top 10 highest expression genes in E. coli.

CIP-T0 by smRandom-seq		Species-mixture by smRandom-seq		Public bulk RNA-seq	
Gene ID	Count proportion	Gene ID	Count proportion	Gene ID	Count proportion
BW25113_RS19530	0.1785308	BW25113_RS19995	0.145887	BW25113_RS19530	0.30111
BW25113_RS01010	0.149809035	BW25113_RS19530	0.066561	BW25113_RS19995	0.211552
BW25113_RS23140	0.101970708	BW25113_RS01025	0.060744	BW25113_RS01025	0.211342
BW25113_RS13535	0.073189522	BW25113_RS19520	0.043218	BW25113_RS13535	0.064297
BW25113_RS01025	0.065432498	BW25113_RS13535	0.042552	BW25113_RS19520	0.058034
BW25113_RS23125	0.055954742	BW25113_RS04685	0.012936	BW25113_RS19980	0.039734
BW25113_RS19995	0.041493869	BW25113_RS17310	0.012513	BW25113_RS17000	0.038523
BW25113_RS13525	0.036868505	BW25113_RS01010	0.012182	BW25113_RS16985	0.029103
BW25113_RS22750	0.030024506	BW25113_RS09310	0.011086	BW25113_RS01010	0.024043
BW25113_RS16985	0.027230109	BW25113_RS00570	0.009709	BW25113_RS23125	0.00807

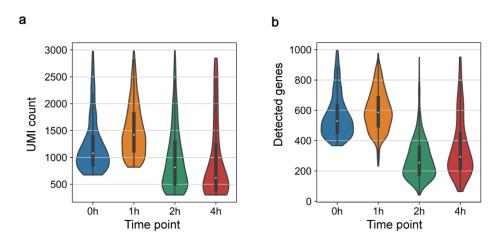
Supplementary Figure 11. Gene expression of CIP-treated E. coli samples.

a, b, Violin plots of the distribution of UMI count (**a**) and detected genes (**b**) of CIP-treated *E. coli* samples. 0h n= 771 cells, CIP 1h n= 850 cells, CIP 2h n= 950 cells, CIP 4h n= 949 cells. Data in the violin plot in **Supplementary fig. 11a, b** corresponded to the median of the distribution (white dot in the center), the interquartile range of the data (the thicker black bar), and the data that extends to 1.5 times the interquartile range (the thinner black bar). **c**, The top 10 highest expression genes sorted by the count proportion in *E. coli* datasets of the CIP-T0 sample, the species mixture by smRandom-seq and the public bulk RNA-seq dataset of the BW25113 Wild Type_ECWT1_1 sample².

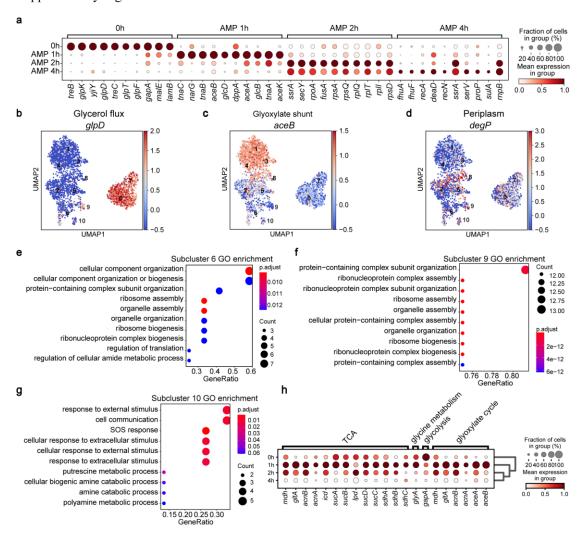


Supplementary Figure 12. Growth curves and morphology of E. coli upon AMP.

a, OD600 curves of *E. coli* exposure to 0 and 5 μ g/mL AMP. Source data are provided as a Source Data file. **b**, Image of *E. coli* samples treated with 5 μ g/mL AMP after 0, 1, 2, and 4 hours. Scale bar: 10 μ m. n= 4 independent experiments.

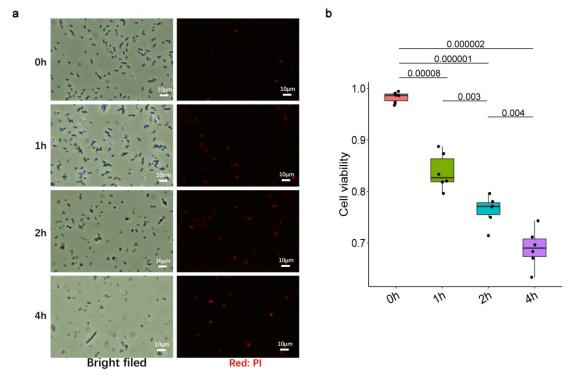


Supplementary Figure 13. UMI count (a) and detected genes (b) of AMP-treated *E. coli* **samples.** 0h n= 831 cells, AMP 1h n= 1094 cells, AMP 2h n= 716 cells, AMP 4h n= 216 cells. Data in the violin plot in **Supplementary fig. 13a, b** corresponded to the median of the distribution (white dot in the center), the interquartile range of the data (the thicker black bar), and the data that extends to 1.5 times the interquartile range (the thinner black bar).



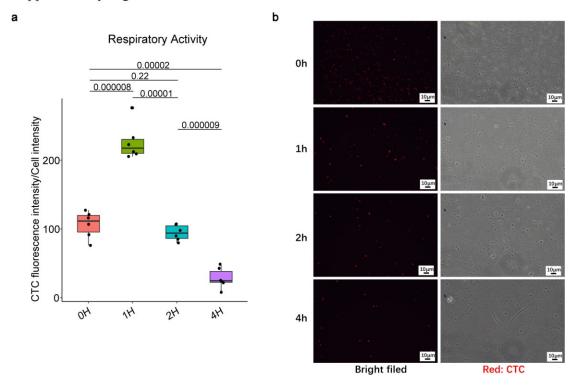
Supplementary Figure 14. Gene expression patterns of E. coli upon 12.5 μg/mL AMP.

a, Top 10 DEGs among groups of *E. coli* samples upon 12.5 μg/mL AMP. **b**, Expression of glycerol flux related marker genes of subclusters 0 and 7 (*glpD*) on the t-SNE plot. **c**, Expression of glyoxylates shunt related marker genes of subclusters 1, 3 and 4 (*aceB*) on the t-SNE plot. **d**, Expression of periplasm related marker genes of subclusters 8 (*degP*) on the t-SNE plot. **e-g**, Bubble plots of GO enrichment analysis of top 20 DEGs in subclusters 6 (**e**), 9 (**f**) and 10 (**g**) of *E. coli* samples upon high concentration of AMP. The cutoff value of p-value of GO enrichment was 0.05 and the cutoff value of q-value of GO enrichment was 0.05. **h**, Expression of genes involved in different metabolic pathways in different groups. TCA: tricarboxylic acid cycle. Source data are provided as a Source Data file.



Supplementary Figure 15. Cell viability analysis of *E. coli* upon 12.5 μg/mL AMP.

a, Cell viability analysis by combined propidium iodide (PI) staining. Representative fluorescence and bright filed images of *E. coli* samples at 0h, 1h, 2h, 4h time point after 12.5 μ g/mL AMP treatment. Red color (PI) showed positive staining of dead cells. Scale bar showed 10 μ m. n= 4 independent experiments. b, Graphic representation of statistical result of cell viability. n= 6 biologically independent samples. Data in the box plot corresponded to the first quartiles (lower hinges), median (center), and third quartiles (upper hinges). Whisker is defined as 1.5* IQR (interquartile range) from the hinge. Two-tailed Student's *t* test were used for calculation of *p* values. Statistical significance was established at 95% confidence level (*p* values < 0.05). Source data are provided as a Source Data file.



Supplementary Figure 16. Cell respiratory activity analysis of *E. coli* upon 12.5 μg/mL AMP.

a, Cell respiratory activity analysis by the 5-cyano-2,3-ditolyltetrazolium chloride (CTC) dye staining in combination with fluorescence microplate reader. Graphic representation of statistical result of normalized CTC fluorescence intensity (CTC fluorescence intensity/Cell count) of *E. coli* samples at 0h, 1h, 2h, 4h time point after 12.5 μ g/mL AMP treatment. n= 6 biologically independent samples. Data in the box plot corresponded to the first quartiles (lower hinges), median (center), and third quartiles (upper hinges). Whisker is defined as 1.5* IQR (interquartile range) from the hinge. Two-tailed Student's *t* test were used for calculation of *p* values. Statistical significance was established at 95% confidence level (*p* values < 0.05). **b**, Representative fluorescence and bright filed images of *E. coli* samples at 0h, 1h, 2h, 4h time point after 12.5 μ g/mL AMP treatment. Red color (CTC) showed cells with high cell respiratory activity. Scale bar showed 10 μ m. n= 4 independent experiments. Source data are provided as a Source Data file.

Supplementary Tables

Supplementary Table 1: Primers

Primer list	Sequence
Barcoded beads primer	/5Acryd/ATTATATAT U GTG AGT GAT GGT TGA GGA TGT GTG GAGATA [10 bases barcode1] TGGT [10 bases barcode2] GAGA [10 bases barcode3] NNNNNNNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
Random primer	GGAGTTGGAGTGGATGAGT GATG NNNNNNN
PCR-1 primer	GGAGTTGGAGTGGATGAGTGATG
PCR-2 primer	GTG AGT GAT GGT TGA GGA TGT GTG GAG ATA

Supplementary Table 2: Barcode sequences

B	Barcode1_sequences	Barcode2_sequences	Barcode3_sequences
1 (GATGTACATG	ATAGAAACGA	GCTTTTCTTG
2 T	TCAGATTCT	TCCGTGCAAG	CTTTAAGCAT
3 (CGTACTACGC	AATGGGACAC	GAAATATTGT
4 (GCTATTATAG	TATTAGGCGA	GGCGGAAACG
5 C	CAGAAGGAAC	CAATCATCAC	AATGGGTATG
6 A	GACTTTTAA	CAGTTTTACT	CGGCAATTGA
7 (GCAACTGTG	CTTTCTAGAT	CTCGAGGAAT
8 A	AATTCGCGG	TGTGATACCT	CCTATTGTCT
9 (CTTAGACACT	TGTTAACAAC	GAGGCTCCTA
10 A	GCGGACTAT	TTGCTTTCAG	ACTCAAGGGA
11 C	GAGACGCTAC	CAGGAACTAA	ACATCCTGCT
12 C	GTCATCGTGC	TCGGGCGGAT	ATTATGGTAG
13 T	CATTAACAG	ACGTATAGTG	TACCGGATTA
	GCTTCGCTTA	TACTTAAGGC	CAATTTGAAC
15 C	CTGGCAAAGA	CATGCGGGTG	TTTATAGACC
16 T	CGCTTGAAG	TCGCACTCCG	ATCACTCGGT
	TGAACTTTG	GAAACTGAAA	ACTACTAATA
18 C	CAGAGATACG	TACAGGCGCA	CCGGAGGATA
19 (GCTGATAAA	GTTGGCATTT	GACGGGAGGT
	ATTTGTCGC	GTTTTAGGAA	TCTGGAATAG
	TAAGGCAAA	AGATTGGCGG	TGGAGGGAGC
	CAGTATTAT	GAGCGGGTTT	GCTTGGCAAC
	GGATGCTAT	CGGGAAAGGG	GAGGAATCCT
	CGAATTTCG	TTCTTAGGCG	AAGTAAACGG
	ACTATGGCG	AACACAATAA	TTTTGTGGTG
	AGCTTAGTG	TTCCGGCGAG	TCATTGATTG
	TCAGGCATG	TTTGTTTTAT	CAGTATCATC
	GTTGGTGGT	GCAGGTGCAG	CACATGGTGT
	CATGTAGCT	ATAAGGACCT	GTATAGAGTG
	GCGAACAATG	TCAGTTATCC	TCGGGTAATC
	GTTGCTTGC	AATGGAATGT	AATTAAGAGC
	AACAGCAAT	TGTAATTGAT	TCACAGCACA
33 (GTTTGGACTC	TCAGTGGGTT	AATTACCTCT
	CATTATGCTA	GTGGACGTTG	TGTTCATAGA
	GAGTAAGTGA	CTCCTGCAAA	CACGGTGTAG
	GAGCGGTGTA	TTCGGATTCC	GCCGAGGTGT
_	CGTAAGTCGA	AATCGGGTCA	ATGAGGCTTT
38 T	GGCACGTTT	AGTCGCCGAC	TGATGGCGGA
39 (GAAATCTTTT	CATAAGAAGT	CGCTCCTGGT
	ACAACGCAA	AGGCAGGAAA	TCAAGGCCTG
41 A	TGTTGTGGC	CTTATCCTTG	ATTAAAGGCT
	AACTCATTC	GGTCGAAGGG	CGGAAAGAGT
	CGACTATAT	CAAGAGGTGA	ACAATAGGCG
44 T	TATCAGTCTG	CTACTGGTGG	TCCGACCGTC
	CTATATGTAG	TATGTTCATC	GAAGATACAA
	GCAATACGG	AGGCTTGTTC	CTAAAGGACG
	AACTGGTGT	TCGACAGCAG	TATAAAGGAA
	ATTAGGCGT	GTAAATGGGA	ACAGGAGTGA
	CTGTTCAGG	CGGTAATACG	
	CAGTATTGTT	CATTAGTAAA	
	GTACTATAGT	TTAAAGGTTC	
_	CAGACATTT	TAGAAATGGT	
	TCTTTGGAT	CAACAATGGG	
	CTTGATCTAT	AAATATCGAA	1

55 GTATGLAGG GCCTGGTAGT 56 TATTCAACTG AAAGACTCTT 57 TTTTGACGGA TTAACACTTA 58 ATGGTAGAAA GTTAAGGCTA 59 GTGCTCATC GTCACAGGTT 60 GTTATAGTTG TTTGAAAGTG 61 AGGATCTAC GATTGGGTGC 62 TGCATTTGAA TAAGGATGT 63 AATTCGGATA TATGAGATG 64 GGTTGAGAGC GATGATCGAC 65 CAAGGCTCAT TAGTGGAGTC 66 GAATGGAAGC GCTAAACTTC 67 CTCTCTATATG TTTTGATTAA 68 TCTGGATGA ACTGGTTTAGA 69 TTGATCAGTA ATGAGTGAA 69 TTGATCAGTA ATGAGTGAA 71 TACGACTGAA TCTTGGCTCA 72 AAAGTTCCAA CGACAAGATG 73 GGCGGGATT AGCATTTCT 74 CGGCGGATTT AGCATTTCT 75 CGACAAAGTG TACTAAGTAG 76 TGCGTAATTC			
57 TTTTGACGGA TTAACACTTA 58 ATGGTAGAAA GTTAAGGCTA 59 GTGCTCTATC GTCACAGGTT 60 GTTATAGTTG TTTGAAAGTG 61 AGGATCTTAC GATTGGGTGC 62 TGCATTTGAA TAAAGATGT 63 AATTCGGATA TATTCCAATAT 64 GGTTGAGAGC GATGGTGCC 65 CAAGCTCAT TAGTGGAGTC 66 GAATGGAGC GCTAAACTTC 67 CTTCTATATG TTTTGAATAA 68 TCTGGATTAC ACTGTTTAGA 69 TTGATCAGTA ATGTGTATA 60 TTGATCAGTA ATGTGTATA 61 TTGATCAGTA ATGTGGAGTC 62 TGCATTTGAA TATTCCAATAT 63 AATTCGGATA ATGTGGAGTC 64 GAATGGAAGC GCTAAACTTC 65 CAAGCTCAT TAGTGGAGTC 66 GAATGGAAGC GCTAAACTTC 67 CTTCTATATG TTTTGATTAA 68 TCTGGATTAC ACTGTTTAGA 69 TTGATCAGTA ATGAGTGATA 70 ATTTTTGTCG CTATACCTCC 71 TACGACTGAA TCTTGGCTCA 72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCCTGGAG 74 CGGCGGATTT AGCATTTTCT 75 CGACAAAGTG TACTAAGTAG 76 TGCGTAATTC AGATTAAGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATAGGAGGG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGACT ATGGGTCGC 83 TAAGATACC CACGGGTCAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGAAA 86 ATATGTGGA GCAGAAAA 87 ACACAGTATA TCGCCGGTTA 88 ACGCATAGAT ATGGGTCGC 88 TAAGAATGCG TAAATGTAAA 86 ATATGTGGA GCAGTAAAA 87 ACACAGTATA TCGCCGGTTA 88 ACGCATAGAT ATGGGTCGC 90 CTACATCTC CGGGTCATC 91 CTTCTGGTGG CTGCGGAAG 94 GTAAAGACGT TTGATACAAAT 95 TTGCTACAAC CCTAGGTCCC 91 CTTCTGGTGG CTGCGGGAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCCC	55	GTATGTCAGG	GCCTGGTAGT
58 ATGGTAGAAA GTTAAGGCTA 59 GTGCTCTATC GTCACAGGTT 60 GTTATAGTTG TTTGAAAGTG 61 AGGATCTTAC GATTGGGTGC 62 TGCATTGAA TAAAGGATGT 63 AATTCGGATA TATTCAATAT 64 GGTTGAGAGC GATGGGAGC 65 CAAGGCTCAT TAGTGGAGTC 66 GAATGGAAGC GATAGGAGTC 67 CTTCTATATG TTTTGATAA 68 TCTGGATTAC ACTGTTTAGA 69 TTGATCAGTA ATTTTGATTAA 60 TTGATCAGTA ATTTTTGATTAA 61 TAGTGGAGGC GATGGAGTC 62 TGCATTATG TTTTGATTAA 63 TCTGGATTAC ACTGTTTAGA 64 TCTGGATTAC ACTGTTTAGA 65 TCTGCATTAC ACTGTTTAGA 66 TCTGGATTAC ACTGTTTAGA 67 TTGATCAGTA ATGAGTGATA 70 ATTTTTGTCG CTATACCTCC 71 TACGACTGAA TCTTGGCTCA 72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCCTGGAG 74 CGGCGGATTT ACTAAGTAG 75 CGACAAAGTG TACTAAGTAG 76 TGCGTAATTC AGTTTAAGGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATACAGTAGG 79 ACACGGCTCT TATCAGGTC 80 ATTATGACTG GAATGCCAA 81 ACGCATAGAC CACGGGTCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGAC CACGGGTCAA 83 TAAGTAGATA TCGCCGGTAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTAA CGGATACAAA 87 ACACAGTATA TCGCCGGTTA 88 ACACAGTATA TCGCCGGTTA 88 ACACAGTATA TCGCCGGTTA 89 ACTATGCGA GAATGTCT 88 ACACAGTATA TCGCCGGTTA 89 ACTATGCGA GAATGTCT 80 ATTATGACTG GAATGCAAA 81 CACAAGTATA TCGCCGGTTA 82 TGGAAGGAC TAAATGTAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTTGAA GCAGTAGAAA 87 ACACAGTTGTA ATTCTTCAA 88 ACACAGTTAT ATTCTTCAA 89 ACTATCGGAG TAAATGTAAC 90 CTACATGCC CGGGTCATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCAAAA 94 GTAAAGACGT TGAATACAAAT 95 TTGCTACAAC CCTAGGTCAC	56	TATTCAACTG	AAAGACTCTT
59 GTGCTCTATC GTCACAGGTT 60 GTTATAGTTG TTTGAAAGTG 61 AGGATCTTAC GATTGGGTGC 62 TGCATTTGAA TAAAGGATGT 63 AATTCGGATA TATTCAATAT 64 GGTTGAGAGC GATGATCGAC 65 CAAGGCTCAT TAGTGGAGTC 66 GAATGGAAGC GCTAAACTTC 67 CTTCTATATG TTTTGATTAA 68 TCTGGATTAC ACTGTTTAGA 69 TTGATCAGTA ATGAGTGATA 70 ATTTTTGTCG CTATACCTCC 71 TACGACTGAA TCTTGGCTCA 72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCCTGGAG 74 CGGCGGATTT AGCATTTCT 75 CGACAAAGTG TACTAAGTAG 76 TGCGTAATTC AGTTTAAGGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATAGGTAGG 79 ACACGCTCT ATAGGTAGGG 79 ACACGGCTT </td <td>57</td> <td>TTTTGACGGA</td> <td>TTAACACTTA</td>	57	TTTTGACGGA	TTAACACTTA
60 GTTATAGTTG TTTGAAAGTG 61 AGGATCTTAC GATTGGGTGC 62 TGCATTTGAA TAAAGGATGT 63 AATTCGGATA TAATCAATAT 64 GGTTGAGAGC GATGATCGAC 65 CAAGGCTCAT TAGTGGAGTC 66 GAATGGAAGC GCTAAACTTC 67 CTTCTATATG TTTTGATTAA 68 TCTGGATTAC ACTGTTTAGA 69 TTGATCAGTA ATGAGTGATA 70 ATTTTTGCG CTATACCTCC 71 TACGACTGAA TCTTGGCTCA 72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCTGGAG 74 CGGCGGATTT AGCATTTCT 75 CGACAAAGTG TACTAAGGA 76 TGCGTAATTC AGTTTAAGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATGAGGAGC 79 ACACGGCTCT TATCAGGGG 79 ACACGGCTCT TATCAGGGG 80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGAC CACGGGTCAA 83 TAAGTAGAA CGGATAACAAA 84 CACAAGTATA CGGATAACAAAA 85 TAAGAATGCG TATCAGAAAA 86 ATATGTGGA CACGGGTCAA 87 ACACAGGCTC ATGAGCAAAA 88 CACAAGTAAA CGGATACAAA 89 ACTAGCGC TAAAGTAAAAAAAAAAAAAAAAAAAAAAAA	58	ATGGTAGAAA	GTTAAGGCTA
61 AGGATCTTAC GATTGGGTGC 62 TGCATTTGAA TAAAGGATGT 63 AATTCGGATA TATTCAATAT 64 GGTTGAGAGC GATGATCGAC 65 CAAGGCTCAT TAGTGGAGC 66 GAATGGAAGC GCTAAACTTC 67 CTTCTATATG TTTTGATTAA 68 TCTGGATTAC ACTGTTTAGA 69 TTGATCAGTA ATGAGTGATA 70 ATTTTGTCG CTATACCCC 71 TACGACTGAA TCTTGGCTCA 72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCCTGGAG 74 CGGCGGATTT AGCATTTTCT 75 CGACAAAGTG TACTAACTACG 76 TGCGTAATC AGTTTAAGA 77 CAAATGTCGT AAACTGGCT 78 CAGCTAGCTT ATGAGTAGGA 79 ACACGGCTCT TATCAGGGG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGACT ATGGGCCC 83 TAAGTAGAC CACGGGTCAA 84 CACAAGTAA CGGATACAAA 85 TAAGAATGCG TACTAGACCC 86 TAGGATAAA CGGATACAAA 87 CACAAGTAAA CGGATACAAA 88 CACAAGTAAA CGACAGATG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCCCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGACT ATGGGTCCC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGGA GCAGTGAAA 87 ACACAGTGTG CGAGTACAAA 88 CACAAGTATA TCGCCGGTTA 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAACC 90 CTACATGCCC 90 CTACATGCCC CGGGTCATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCA 93 CAGCGAAGGA GTAACCCAA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC	59	GTGCTCTATC	GTCACAGGTT
61 AGGATCTTAC GATTGGGTGC 62 TGCATTTGAA TAAAGGATGT 63 AATTCGGATA TATTCAATAT 64 GGTTGAGAGC GATGATCGAC 65 CAAGGCTCAT TAGTGGAGC 66 GAATGGAAGC GCTAAACTTC 67 CTTCTATATG TTTTGATTAA 68 TCTGGATTAC ACTGTTTAGA 69 TTGATCAGTA ATGAGTGATA 70 ATTTTGTCG CTATACCCC 71 TACGACTGAA TCTTGGCTCA 72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCCTGGAG 74 CGGCGGATTT AGCATTTTCT 75 CGACAAAGTG TACTAACTACG 76 TGCGTAATC AGTTTAAGA 77 CAAATGTCGT AAACTGGCT 78 CAGCTAGCTT ATGAGTAGGA 79 ACACGGCTCT TATCAGGGG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGACT ATGGGCCC 83 TAAGTAGAC CACGGGTCAA 84 CACAAGTAA CGGATACAAA 85 TAAGAATGCG TACTAGACCC 86 TAGGATAAA CGGATACAAA 87 CACAAGTAAA CGGATACAAA 88 CACAAGTAAA CGACAGATG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCCCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGACT ATGGGTCCC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGGA GCAGTGAAA 87 ACACAGTGTG CGAGTACAAA 88 CACAAGTATA TCGCCGGTTA 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAACC 90 CTACATGCCC 90 CTACATGCCC CGGGTCATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCA 93 CAGCGAAGGA GTAACCCAA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC	60	GTTATAGTTG	TTTGAAAGTG
63 AATTCGGATA TATTCAATAT 64 GGTTGAGAGC GATGATCGAC 65 CAAGGCTCAT TAGTGGAGTC 66 GAATGGAAGC GCTAAACTTC 67 CTTCTATATG TTTTGATTAA 68 TCTGGATTAC ACTGTTTAGA 69 TTGATCAGTA ATGAGTGATA 70 ATTTTGTCG CTATACCTCC 71 TACGACTGAA TCTTGGCTCA 72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCCTGGAG 74 CGGCGGATTT AGCATTTCT 75 CGACAAAGTG TACTAAGTAGAG 76 TGCGTAATTC AGTTAAGGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCGACTT ATACGAGGG 79 ACACGGCTCT TATCAGGGCC 80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGACC CACGGGTCAA 83 TAAGTAGAC CACGGGTCA 84 CACAGGTCT ATGGGTCGC 85 TAAGTAGAA 86 ATATGTGTG GAATGCCC 87 TAAGTAGAA 87 ACACAGGTC TACGAGAA 88 TAAGTAGATA CGGATACAAA 89 ACACGGTTA 88 AGCACTGTAA 87 ACACAGTTC 88 AGCACAGTAC 88 TAAGTAGATA TCGCCGATA 89 ACACGGCTC 80 CACGGGTCA 81 TAAGTAGATA CGGATACAAA 84 CACAGTATA TCGCCGGTTA 85 TAAGTAGATA TCGCCGATA 86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTC CGGGTCAA 88 TAAGTAGATA TCGCCGGTTA 89 ACTATGCGCA GCGGTCAA 80 TATTGCTGA GCAGTGGAAA 81 TAAGTAGATA TCGCCGGTTA 82 TAGGAATGCC TAAATGTAAA 83 TAAGTAGATA TCGCCGGTTA 84 TATGTGTGA GCAGTGGAAA 85 TAAGTAGACC CACGGTCC 86 ATTATGCTGA GCAGTGGAAA 87 ACACAGTTT ATTCTCAA 88 ACCACTGTAT TATTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCTATC 91 CTTCTGGTGG CTGCGGGGATG 92 GGATGCTCGA GCGGTGCTCT 93 CAGCGAAGGA GTAAACGCAGA 94 GTAAAGACGT TGATACACAAT	61	AGGATCTTAC	GATTGGGTGC
64 GGTTGAGAGC 65 CAAGGCTCAT 66 GAATGGAAGC 66 GAATGGAAGC 67 CTCTATATG 67 CTTCTATATG 67 TTTTGATTAA 68 TCTGGATTAC 69 TTGATCAGTA 70 ATTTTTGTCG 71 TACGACTCA 71 TACGACTGAA 72 AAAGTTGCAA 73 GGCGTGCTTT 74 CGGCGGATTT 75 CGACAAAGTT 76 TGCGTAATTC 77 TACGACTGA 77 CAAATGTCGT 78 CAGCGAGATG 79 ACCGGCTCT 78 CAGCGAGCT 79 ACAGGCTCT 80 ATTATGACTG 80 ATTATGACGC 81 ACCGGCGAA 81 ACGCATAGAC 82 TGGAAGAC 82 TGGAAGAC 83 TAAGAAGAC 84 CACAAGTA 85 TAAGAAGAC 86 ATATGTGGA 87 ACACAGTGT 88 AGCACTGATA 88 ACACAGTTA 88 ACACAGTG 88 ATATGAAGA 87 ACACAGTGC 88 ATATGAAA 88 ACACAGTAT 88 AGCACTGATA 89 ACTATGCGA 89 ACTATGCGA 80 CAGGGCTC 80 ATTATAAA 81 ACGCATAGAC 82 TGGAAGAC 83 TAAGAAGAC 84 CACAGGTTA 85 TAAGAATA 86 ATATGTGTA 87 ACACAGTGT 88 AGCACTGATA 89 ACTATGCG 90 CTACATGCC 90 CTACATGCC 91 CTCCTGGAA 91 GTAAAGCAA 94 GTAAAGACA 95 TTGCTACAAC CCTAGGTCAC CCTAGGTCAC CCTAGGTCAC CCTAGGTCAC CCTAGGTCC CCTAGGTCC CCTAGACAC CCCAGGTCAA CCTAGAGAC CACGGGTCAA CCTAGACC CACGGTTAA CCGCATACAA CCGCTTAA CCGCTTAA CCGCTTAA CCGCGTTAA CCGCGTTAA CCGCGGTTAA CCGCGGTTAA CCGCGGTTAA CCGCGGTCAA CCGCGTCAA CCGCGTCAA CCGCGTCAA CCGCGTCAA CCGCGTCAA CCCTAGGTCC CCGGGTCAAC CCTAGGTCC CCGGGTCAT CCGCGGATAC CCCTAGGTCC CCGGGTCAAC CCCTAGGTCC CCCTAGGTCC CCCTAGGTCAC	62	TGCATTTGAA	TAAAGGATGT
65 CAAGGCTCAT TAGTGGAGTC 66 GAATGGAAGC GCTAAACTTC 67 CTTCTATATG TTTTGATTAA 68 TCTGGATTAC ACTGTTAGA 69 TTGATCAGTA ATGAGTGATA 70 ATTTTTGTCG CTATACCTCC 71 TACGACTGAA TCTTGGCTCA 72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCTGGAG 74 CGGCGGATTT AGCATTTCT 75 CGACAAAGTG TACTAAGTAG 76 TGCGTAATTC AGTTTAAGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATAGGAGG 79 ACACGGCTCT TATCAGGGC 80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGAAC CACGGGTCAA 82 TGGAAGACT ATGGGTCC 83 TAAGTAGAC CACGGGTCAA 84 CACAAGTATA CCGGATACAA 85 TAAGTAGATA CCGATACAAA 86 ATATCTGTG GAATGCGCC 87 TAAATGTCAA 88 CACCATGATA TCCGCGGTTA 88 AGCACTAGAT ATGGGTCCC 89 TAAATGTAAA 89 CACAGGTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCC 81 TAAGTAGATA CCGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATCTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGAC 88 AGCACTGTAT ATTCTTCAA 89 ACTATCGGAG TAAATGTAAA 89 ACTATCGGAG TGTTTGATCC 90 CTACATGCTC CGGGTCTATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCGA GCGGTCCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACCT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC	63	AATTCGGATA	TATTCAATAT
66 GAATGGAAGC 67 CTTCTATATG 68 TCTGGATTAC 68 TCTGGATTAC 69 TTGATCAGTA ATGATCAGTA ATGATCAGTA ATGATCAGTA ATGATCAGTA ATGATCAGTA ATGATCAGTA ATTTTTGCG CTATACCTCC 71 TACGACTGAA TCTTGGCTCA 72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCCTGGAG 74 CGGCGGATTT AGCATTTCTT 75 CGACAAAGTG TACTAAGTAG ATTTAAGGA ATGATTAAGGA ATGATTAAGGA ATGATTC AAATGTCGT AAACTGGCCT AAACTGGCT AAACTGGCT AAACTGGCT AAACTGGCT AAACTGGCT ATACAGGGG ATTATGACTG AAATGAGGG ATTATGACTG AAATGAGGA ATTATGACTG AAATGAGGA ATTATGACTG AAATGAGAA ATTATGACTG AAATGAAGA ATTATGACTG AAATGAAGA ATTATGACTG AAATGAAAA ATTAGACTG AATGAAGAAA ATGACGCAA ATGACGCAA ATGACGCAA ATGACGAAAA ATGACGAAA ATGACGAAA ATGACTAAAA ATGACTAAAA ATGACTAAAA ATGACTAAAA ATAGATAGATA ATGACTAAAA ATAGATAGATA ATTCTTCAA ATAGATTGAAA ATATGATGAAA ATTTCTTCAA ATTTCTCAA ATTTCTCAA ATTTCTCAA ATTTCTCAA ATTTCTCAA ATTTCTCAA ATTTCTCAA ATTTCTCAA ATTTCTCA		GGTTGAGAGC	GATGATCGAC
67 CTTCTATATG TTTTGATTAA 68 TCTGGATTAC ACTGTTTAGA 69 TTGATCAGTA ATGAGTGATA 70 ATTTTTGTCG CTATACCTCC 71 TACGACTGAA TCTTGGCTCA 72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCCTGGAG 74 CGGCGGATTT AGCATTTCT 75 CGACAAAGTG TACTAAGGA 76 TGCGTAATTC AGTTTAAGGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATACGACGA 79 ACACGGCTCT TATCAGGGG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCCCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGAC CACGGGTCAA 82 TGGAAGAC CACGGGTCAA 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTAA 87 ACACAGTGTA ATAGGTAAAA 88 ACACTGGCC TAAATGTAAA 89 ACTATGCGA GCAGTGAAAA 80 ATTTCTTCAA 81 ACCATGCCC CGGGTCAA 82 TGCACGGTCT TATCCCGCTTAACCCC 83 TAAGAATGCG TAAATGTAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTAA GCAGTAGACA 87 ACACAGTGTA ATTTCTTCAA 88 ACCACTGTAT ATTTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCAA GCGGTCCT 93 CAGCGAAGAA GTAACGCAGA 94 GTAAAGACCT TGAACGCAGA 94 GTAAAGACCT TGAACACAAT 95 TTGCTACAAC CCCTAGGTCAC	65	CAAGGCTCAT	TAGTGGAGTC
68 TCTGGATTAC ACTGTTTAGA 69 TTGATCAGTA ATGAGTGATA 70 ATTTTTGTCG CTATACCTCC 71 TACGACTGAA TCTTGGCTCA 72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCCTGGAG 74 CGGCGGATTT AGCATTTCT 75 CGACAAAGTG TACTAAGTAG 76 TGCGTAATTC AGTTTAAGGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATAGGTAGGG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGAGA CACGGCCA 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAAA 86 ATATGTGTAG GCAGTGGAAA 87 ACACAGTGTG CGAGTGGAAA 88 ACCCTGTAT ATAGGTAGAA 89 ACTATGCGA GCAGTGGAAA 80 ATTTCTTCAA 81 ACGCATAGAC CACGGTCAA 82 TGGAAGACT ATGCCGCTTA 83 TAAGTAGATA TCGCCGGTTA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAAA 86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCAA 94 GTAAAAGACGT TGATACACAAT 95 TTGCTACAAC CCTAGGTCAC	66	GAATGGAAGC	GCTAAACTTC
69 TTGATCAGTA ATGAGTGATA 70 ATTTTTGTCG CTATACCTCC 71 TACGACTGAA TCTTGGCTCA 72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCCTGGAG 74 CGGCGGATTT AGCATTTCT 75 CGACAAAGTG TACTAAGTAG 76 TGCGTAATTC AGTATAAGGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATACGGAGG 79 ACACGGCTCT ATACAGGGG 80 ATTATGACTG GAATGCCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGAC ATGGACCAA 82 TGGAAGGACT ATGGGTCGC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGGA GCAGTGAAA 87 ACACAGTATA TCGCCGGTTA 88 AGCACTGTG TAATGAAAA 89 ACTATGCGA TAATGTAAA 80 ATTTTGACAGGTC 81 TAATGTTGAAA 82 TGGAAGAC CACGGGTCAC 83 TAAGTAGATA TCGCCGGTTA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCTATC 91 CTCTCGGTGG CTGCGGGATG 92 GGATGCTCGA GCGGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACCT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC	67	CTTCTATATG	TTTTGATTAA
70 ATTTTTGTCG CTATACCTCC 71 TACGACTGAA TCTTGGCTCA 72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCCTGGAG 74 CGGCGGATTT AGCATTTCT 75 CGACAAAGTG TACTAAGTAG 76 TGCGTAATTC AGTTTAAGGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATAGGTAGG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGAC ATGGGCCC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTGA GCAGTGAAAA 87 ACACAGTGTG CGAGTGAAA 88 AGCACTGTAT ATTCTCAA 89 ACTATGCGAG TGATTGAAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCTTC 91 CTTCTGGTGG CTGCCGGGAAG 92 GGATGCTCAA 94 GTAAAGACT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC	68	TCTGGATTAC	ACTGTTTAGA
71 TACGACTGAA TCTTGGCTCA 72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCCTGGAG 74 CGGCGGATTT AGCATTTTCT 75 CGACAAAGTG TACTAAGTAG 76 TGCGTAATTC AGTTTAAGGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATACGAGGG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGACT ATGGGTCGCC 83 TAAGATAGAAA CGGATACAAA 84 CACAAGTATA CGGATACAAA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTGA GCAGTGAAA 87 ACACAGTGTG CAAATGTAAA 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGA GCAGTAGTC 90 CTACATGCTC CGGGTCATC 91 CTCTCTGGTGG CTGCCGGAGAG 94 GTAAAGAGAGA GTAACAAAT 95 TTGCTACAAC CTCAAGACA 94 GTAAAGAGAG 94 GTAAAGAGAG 95 TTGCTACAAC CCCTAGGTCAC			
72 AAAGTTGCAA CGACAGAATG 73 GGCGTGCTTT CCTCCTGGAG 74 CGGCGGATTT AGCATTTCT 75 CGACAAAGTG TACTAAGTAG 76 TGCGTAATTC AGTTAAGGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATAGGTAGGG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGACT ATGGGTCGCC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGATGCG TAAATGTAAA 86 ATATGTGTG GAATGCAAA 87 ACACAGTGTG CGAGTGAAA 88 ACACAGTATA TCGCCGGTTA 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGA GCGGTCAA 89 ACTATGCGA GCGGTCAA 89 ACTATGCGA GCGGTCAA 89 ACTATGCGA GCGGTCAC 90 CTACATGCTC CGGGTCTATC 91 CTTCTGGTGG CTGCCGGTTA 92 GGATGCTCGA GCGGTCCT 93 CAGCGAAGGA GTAACAAAT 94 GTAAAGACC CCTAGGTCAC	70	ATTTTTGTCG	
73 GGCGTGCTTT CCTCCTGGAG 74 CGGCGGATTT AGCATTTCT 75 CGACAAAGTG TACTAAGTAG 76 TGCGTAATTC AGTTTAAGGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATAGGTAGG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGACT ATGGGTCGCC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGGA GCAGTGGAA 87 ACACAGTGTG CGAGTGAAA 88 AGCACTGATA ATTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCATC 91 CTTCTGGTGG CTGCCGGTCT 93 CAGCGAAGGA GTAACGAAA 94 GTAAAGACT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC	71	TACGACTGAA	TCTTGGCTCA
74 CGGCGGATTT AGCATTTCT 75 CGACAAAGTG TACTAAGTAG 76 TGCGTAATTC AGTTTAAGGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATAGGTAGG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGACT ATGGGTCGCC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGGA GCAGTGGAAA 87 ACACAGTGTG CAGTGGAAA 88 AGCACTGTAT ATTCTCAA 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCTC 91 CTTCTGGTGG CTGCGGGAA 94 GTAAAGACT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC			
75 CGACAAAGTG TACTAAGTAG 76 TGCGTAATTC AGTTTAAGGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATAGGTAGGG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGACT ATGGGTCGCC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCATC 91 CTTCTGGTGG CTGCGGGAG 92 GGATGCTCGA GCGGTGCTC 93 CAGCGAAGGA 94 GTAAAGACT TGATACAAAT 95 TTGCTACAAC CCCTAGGTCAC	73		
76 TGCGTAATTC AGTTTAAGGA 77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATAGGTAGGG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGACT ATGGGTCGCC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCATC 91 CTTCTGGTGG CTGCGGGTC 92 GGATGCTCGA GCGGTGCTCT 93 CAGCGAAGGA 94 GTAAAGACT TGATACAAAT 95 TTGCTACAAC CCCTAGGTCAC	74	CGGCGGATTT	AGCATTTTCT
77 CAAATGTCGT AAACTGGCCT 78 CAGCTAGCTT ATAGGTAGGG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGACT ATGGGTCGCC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCAG 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC	75	CGACAAAGTG	TACTAAGTAG
78 CAGCTAGCTT ATAGGTAGGG 79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGACT ATGGGTCGCC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCTATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCAA 93 CAGCGAAGGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC	76		
79 ACACGGCTCT TATCAGGGTC 80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGACT ATGGGTCGCC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCATC 91 CTTCTGGTGG CTGCGGATG 92 GGATGCTCAA GCGGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCCTAGGTCAC			
80 ATTATGACTG GAATGCGCAA 81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGACT ATGGGTCGCC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCAA GCGGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCCTAGGTCAC			ATAGGTAGGG
81 ACGCATAGAC CACGGGTCAA 82 TGGAAGGACT ATGGGTCGCC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCTATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCGA GCGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC	79	ACACGGCTCT	TATCAGGGTC
82 TGGAAGGACT ATGGGTCGCC 83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGAG TGTTGATCC 90 CTACATGCTC CGGGTCTATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCGA GCGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC	80		
83 TAAGTAGATA CGGATACAAA 84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCTATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCGA GCGGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC			
84 CACAAGTATA TCGCCGGTTA 85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCTATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCGA GCGGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC	82		
85 TAAGAATGCG TAAATGTAAA 86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCTATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCGA GCGGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC		TAAGTAGATA	CGGATACAAA
86 ATATGTGTGA GCAGTGGAAA 87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCTATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCGA GCGGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC			TCGCCGGTTA
87 ACACAGTGTG CGAGTAGTCT 88 AGCACTGTAT ATTTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCTATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCGA GCGGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC	85	TAAGAATGCG	TAAATGTAAA
88 AGCACTGTAT ATTTCTTCAA 89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCTATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCGA GCGGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC			
89 ACTATGCGAG TGTTTGATCC 90 CTACATGCTC CGGGTCTATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCGA GCGGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC			
90 CTACATGCTC CGGGTCTATC 91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCGA GCGGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC		1	
91 CTTCTGGTGG CTGCGGGATG 92 GGATGCTCGA GCGGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC			
92 GGATGCTCGA GCGGTGCTCT 93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC			
93 CAGCGAAGGA GTAACGCAGA 94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC			
94 GTAAAGACGT TGATACAAAT 95 TTGCTACAAC CCTAGGTCAC			
95 TTGCTACAAC CCTAGGTCAC			
96 TGGAAACGGT TCTTTAGCGG			
	96	TGGAAACGGT	TCTTTAGCGG

Supplementary References

- 1. Wang B, K.F., Hamoen LW Induction of the CtsR regulon improves Xylanase production in Bacillus subtilis. *Preprint from Research Square* (2023).
- 2. Ojo, O., Scott, D., Iwalokun, B., Odetoyin, B. & Grove, A. Transcriptome RNA Sequencing Data Set of Differential Gene Expression in Escherichia coli BW25113 Wild-Type and slyA Mutant Strains. *Microbiol Resour Announc* **10** (2021).