OxyR and SoxRS regulon DESCRIPTION DATA FILE

1. GENERAL INFORMATION.

Title: OxyR and SoxRS regulons

Description of the dataset:

OxyR and SoxRS dataset from ChIP-exo and RNA-seq analysis.

Confidence:

ChIP analysis and statistical validation of TFBSs (CHIP-SV) and Mapping of signal intensities by RNA-seq (MSI).

Reference:

Genome-wide Reconstruction of OxyR and SoxRS Transcriptional Regulatory Networks under Oxidative Stress in Escherichia coli K-12 MG1655. Seo SW, Kim D, Szubin R, Palsson BO. Cell Rep. 2015 Aug 25;12(8):1289-99. doi: 10.1016/j.celrep.2015.07.043. Epub 2015 Aug 13. PMID: 26279566

Citation:

Dataset provided and maintained by RegulonDB (PUBMED: #18158297) from the original source published in*:* (PUBMED: #26258987)

Contact person for this dataset:

Questions concerning the content of the data set that are raised by users of RegulonDB will be forwarded to this person. We would appreciate receiving a copy of the response to the user, so we can keep track of taking care of user requests.
Person: RegulonDB staff
Email address: regulondb@ccg.unam.mx

Licensing:

See the license of RegulonDB in:

http://regulondb.ccg.unam.mx/LicenseRegulonDBWithoutSing.jsp

2. DATASET DESCRIPTION.

Summary:

OxyR and SoxRS regulons were determined under oxidative stress (paraquat treatment) based on ChIp-exo and RNA-seq analysis. These consist of 68 genes in 51 transcription units (OxyR, 38 genes in 28 TUs; SoxR, 11 genes in 10 TUs; and SoxS, 34 genes in 25 TUs; several genes are co- regulated). Among them, 48 genes in 36 TUs (OxyR, 26 genes in 17 TUs; SoxR, 8 genes in 7 TUs; and SoxS, 19 genes in 16 TUs) showed more than 2-fold changes in expression level under single-TF- knockout conditions. Their roles include direct activation of amino acid biosynthesis, cell wall synthesis, and divalent metal ion transport.

28, 10, and 25 binding sites for OxyR, SoxR, and SoxS were identified respectively, under oxidative stress by ChIP-exo experiments. Their target genes were classified into 13 functional cartegories. Except for a few cases where Sigma 38 (RpoS) was involved, most of the promoters were Sigma 70 (RpoD) dependent.

Experiment:

Bacterial Strains, Media, and Growth Conditions

All strains used are *E. coli* K-12 MG1655 and its derivatives. Strains harboring OxyR-8myc, SoxR-8myc, and SoxS-8myc: Cho, B.K., Knight, E.M., and Palsson, B.O. (2006). PCR-based tandem epitope tagging system for *Escherichia coli* genome engineering. Biotechniques 40, 67–72. Deletion mutants (DoxyR, DsoxR, and DsoxS): Datta, S., Costantino, N., and Court, D.L. (2006). A set of recombineering plasmids for gram-negative bacteria. Gene 379, 109–115. Glycerol stocks of strains were inoculated into fresh 70 ml M9 minimal medium (47.8 mM Na2HPO4, 22 mM KH2PO4, 8.6 mM NaCl, 18.7 mM NH4Cl, 2 mM MgSO4, and 0.1 mM CaCl2) supplemented with 0.2% (w/v) glucose in a 500-ml flask and cultured overnight at 37ºC at 250 rpm. To create oxidative stress, the overnight cultures were inoculated to an optical density 600 (OD600) = 0.01 into the fresh 70 ml of M9 minimal medium in a 500-ml flask supplemented with 250 mM paraquat (PQ) at OD600 = 0.3 ± 0.03 and incubated for 20 min with stirring.

ChIP-Exo

To identify OxyR-, SoxR-, and SoxS-binding maps in vivo, we isolated the DNA bound to each TF from formaldehyde cross-linked *E. coli* cells by chromatin immunoprecipitation (ChIP) with the antibodies that specifically recognizes myc tag (9E10, Santa Cruz Biotechnology) and Dynabeads Pan Mouse IgG magnetic beads (Invitrogen) followed by stringent washing steps: Cho, B.K., Barrett, C.L., Knight, E.M., Park, Y.S., and Palsson, B.O. (2008a). Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. Proc. Natl. Acad. Sci. USA 105, 19462–19467. Based on Rhee, H. S. & Pugh, B. F. ChIP-exo method for identifying genomic location of DNA-binding proteins with near-

single-nucleotide accuracy. Curr. Protoc. Mol. Biol. Chapter 21, 21–24 (2012), with following modifications as shown in the previous study: Seo, S. W. et al. Deciphering Fur transcriptional regulatory network highlights its complex role beyond iron metabolism in *Escherichia coli*. Nat. Commun. 5, 4910 (2014) and Seo, S.W., Kim, D., O'Brien, E.J., Szubin, R., and Palsson, B.O. (2015). Decoding genome-wide GadEWX-transcriptional regulatory networks reveals multifaceted cellular responses to acid stress in *Escherichia coli.* Nat. Commun. 6, 7970. Prepared DNA libraries were sequenced using MiSeq (Illumina) in accordance with the manufacturer's instructions. ChIP-exo experiments were performed in biological duplicate. Sequence reads generated from ChIP-exo were mapped onto the reference genome (NC_000913.2) using bowtie: Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25.with default options to generate SAM output files. MACE: model based analysis of ChIP-exo. Nucleic Acids Res. 42, e156 (2014). To reduce false-positive peaks Mock-IP was used: were removed as in the previous study: Seo, S. W. et al. Deciphering Fur transcriptional regulatory network highlights its complex role beyond iron metabolism in *Escherichia coli*. Nat. Commun. 5, 4910 (2014).

RNA-Seq Expression Profiling

Based on: Levin, J. Z. et al. Comprehensive comparative analysis of strand-specific RNA sequencing methods. Nat. Methods 7, 709–715 (2010), with following modifications as shown in the previous study: Seo, S. W. et al. Deciphering Fur transcriptional regulatory network highlights its complex role beyond iron metabolism in *Escherichia coli*. Nat. Commun. 5, 4910 (2014) and Seo, S.W., Kim, D., O'Brien, E.J., Szubin, R., and Palsson, B.O. (2015). Decoding genome-wide GadEWX-transcriptional regulatory networks reveals multifaceted cellular responses to acid stress in *Escherichia coli*. Nat. Commun. 6, 7970. The samples were sequenced using MiSeq (Illumina) in accordance with the manufacturer's instructions. RNA-seq experiments were performed in biological duplicate. Sequence reads generated from RNA-seq were mapped onto the reference genome (NC_000913.2) using bowtie (Langmead et al., 2009) with the maximum insert size of 1,000 bp and two maximum mismatches after trim- ming 3 bp at 30 ends. These files were then used for Cufflinks (http://cufflinks. cbcb.umd.edu/) (Trapnell et al., 2010) and Cuffdiff to calculate fragments per kilobase of exon per million fragments (FPKM) and differential expression, with default options and library type of dUTP RNA-seq. From cuffdiff output, genes with differential expression with log2 fold change R1.0 and an FDR % 0.01 were considered as differentially expressed genes.

Motif Search and Analysis

Binding motif analyses were carried ot using the MEME tool from the MEME software suite with default settings (Bailey et al., 2009). Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 37, W202–W208.

COG Functional Enrichment

The OxyR, SoxR, and SoxS regulons were categorized according to their an- notated clusters of orthologous groups (COG) category.

Conservation Analysis of OxyR, SoxR, and SoxS Regulons

Gene annotation of strains and species: (http://theseed.org), and ortholog calculation: RAST (Rapid Annotation using Subsystem Technology) server: Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., et al. (2008). The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9, 75. Conservation levels of oxyR, soxR, soxS, and genes that belong to OxyR, SoxR, and SoxS regulons were calculated from orthologs retained from the RAST output.

Susceptibility Assays under Oxidative Stress

Based on: Minakami, H., and Fridovich, I. (1990). Relationship between growth of *Escherichia coli* and susceptibility to the lethal effect of paraquat. FASEB J. 4, 3239– 3244. Percent survival was calculated as follows: [(CFU/ml at time 2 hr or 4 hr)/ (CFU/ml at time 0)] 3 100. The results presented are averages of triplicate experiments and include the SDs.

More details for each Methods, see: Genome-wide Reconstruction of OxyR and SoxRS Transcriptional Regulatory Networks under Oxidative Stress in *Escherichia coli* K-12 MG1655. Seo SW, Kim D, Szubin R, Palsson BO. Cell Rep. 2015 Aug 25;12(8):1289-99. doi: 10.1016/j.celrep.2015.07.043. Epub 2015 Aug 13. PMID: 26279566

Methods:

Version of programs:

Not apply

Version of datasets:

Not apply

Protocol or algorithm

See Experiment section.

Specificity and sensitivity:

Not apply.

3. COLUMN FORMAT OF DATA FILE BY

Version of E.coli´s Genome: *E. coli* K-12 MG1655 and its derivatives. The version used is not mentioned.

Sequence Identifier:

Suggested column format of the data file:

1) Transcription Factor Name

2) Gene or operon regulated by the TF (regulated gene)

3) Regulatory effect of the TF on the regulated gene (+ activator, - repressor, +- dual, nd not determined)

4) ChIP-exo Start

5) ChIP-exo End

6) Distance to TSS (Transcription Start Site) or C.P. Central position of the TF binding-site

7) Growth condition of the experimental procedures.

8) Evidence that supports the existence of the regulatory interaction

9) PMID Reference(s)

10) Sigma-type

11) S/N ratio. The signal-to-noise (S/N) ratios of the OxyR-, SoxR-, and SoxS-binding peaks and used them as a proxy of the in vivo binding intensity of each binding site.

12) Significance (P-value) of similarity to each box

13) TFBSs in RegulonDB