Evolution of Transcriptional Regulation of Histidine Biosynthesis and Uptake Genes in Gram-positive Bacteria

German A. Ashniev[1], Natalia V. Sernova[1], Alexey E. Shevkoplias[2], Ilia A. Zharov[1], Mikhail S. Gelfand [1], and Dmitry A. Rodionov[1]

[1]Institute for Information Transmission Problems, Moscow, Russia

[2]National Research University Higher School of Economics, Moscow, Russia
rodionov@burnham.org

Abstract. Genome-scale mapping and reconstruction of metabolic pathways and transcriptional regulatory networks in taxonomically diverse microbes is one of the critical tasks of microbial genomics. Here we present a bioinformatic analysis of transcriptional regulatory mechanisms of histidine biosynthetic and transport genes across the reference set of 628 representative genomes from the Firmicutes phylum. By using the comparative genomics approach we have reconstructed the biochemical pathways and transcriptional regulons for histidine metabolism in these genomes and propose a plausible evolutionary scenario for regulation of histidine biosynthesis genes. Orthologs of HisR repressor were identified in 397 Firmicutes genomes, representing four classes (Bacilli, Clostridia, Negativicutes, and Tissierellia), with the only exception of bacteria from the Erysipelotrichia class, where a novel potential RNA attenuator was identified in the leader region of his operon. Histidine attenuator was also identified in a small group of closely-related *Bacillus* species and some *Clostridia* spp. Another class of *cis*-regulatory RNA elements, namely T-boxes, were found to be involved in the transcriptional control of histidine metabolism in bacteria from the Lactobacillales order. We also reconstructed HisR regulons in bacteria from other taxonomic groups from the Actinobacteria, Proteobacteria and Synergistetes phyla. The predicted DNA binding motifs of HisR in various taxonomic groups have a structure of 20-bp palindromes. The reconstructed HisR regulons mostly include the histidine bio-synthetic operons and putative histidine uptake transporters.

Keywords: Histidine Metabolism, Transcriptional Regulation, Gram-Positive Bacteria.

1 Introduction

Histidine is one of the essential amino acids in mammals (1, 2). Prokaryotes, fungi, and plants can synthesize it using a common biosynthesis pathway. Histidine biosynthetic pathway in Bacteria consists of ten biochemical steps. In *Escherichia coli* and other

gamma-proteobacteria, the histidine biosynthesis genes are organized into the his operon, which is regulated via the leader peptide-dependent transcription attenuation mechanism (3). However, the histidine-dependent transcriptional regulation is poorly studied in Gram-positive bacteria from the Firmicutes phylum. In our previous bioinformatic analysis of transcriptional regulation in two Firmicutes genera (*Bacillus* and *Staphylococcus*), we have identified a novel potential regulator of the histidine metabolism, termed HisR that is encoded by the hypothetical gene *yerC* in *B. subtilis* (4,5). The predicted histidine-responsive repressor HisR is homologous to the TrpR family of tryptophan-sensing repressors. The inferred HisR regulons include the histidine biosynthesis his operon and the putative histidine uptake permease *yuiF* in the *Bacillus-Staphylococcus* group (4,5). Another Firmicutes lineage, the Lactobacillales, uses another class of cis-regulatory RNA elements called T-boxes for the transcriptional control of histidine metabolism (6,7).

2 Materials and methods

Histidine biosynthesis metabolic pathways were recosntructed using metabolic subsystems approach implemented in the SEED annotation/analysis tool which combines protein similarity search, positional gene clustering, and phylogenetic profiling of genes (8). Additional analysis was made with Pfam (9) and UniProt/SwissProt (10) databases to verify functional annotations. Transporters were classified according to TCDB database (11). Phylogenetic trees were built with PhyML program (12). The HisR regulons were reconstructed by the RegPredict web server (13) using comparative genomics approach described earlier (14). Multiple sequence alignments were built in MAFFT (15). RNA regulatory motifs were identified using Riboswitch Scanner (16) and Search Rfam (17).

3 Results

Here we present a bioinformatic analysis of transcriptional regulatory mechanisms of histidine biosynthetic and transport genes across the reference set of 628 representative genomes from the Firmicutes phylum. By using the comparative genomics approach we have reconstructed the biochemical pathways and transcriptional regulons for histidine metabolism in these genomes and propose a plausible evolutionary scenario for regulation of histidine biosynthesis genes. Orthologs of HisR repressor were identified in 397 Firmicutes genomes, representing four classes (Bacilli, Clostridia, Negativicutes, and Tissierellia), with the only exception of bacteria from the Erysipelotrichia class, where a novel potential RNA attenuator was identified in the leader region of his operon. A small group of closely-related Bacillus species including *B. cereus*, *B. anthrasis* and *B. thuringiensis* has lost hisR repressor gene and utilizes a novel histidine-specific RNA attenuator for control of the his operon. Histidine attenuator was also identified in some *Clostridia* spp. such as *C. difficile* and related species that lack hisR repressor gene. These findings suggest that the ancestor HisR regulatory system has been replaced by different RNA regulatory systems including T-box and multiple types

of histidine attenuators in certain taxonomic groups of Firmicutes. The reconstructed HisR regulons mostly include the histidine biosynthetic operons and putative histidine uptake transporters.

Using the genome context analysis and regulon reconstruction we observed three major transport systems for histidine uptake in HMP genomes: two distinct ABC-family transport systems (HisXYZ and HisJQMP) and a potential histidine permease from the Na+/H+ antiporter family (YuiF).

We also reconstructed HisR regulons in bacteria from other taxonomic groups from the Actinobacteria, Proteobacteria and Synergistetes phyla, where HisR almost exclusively controls the his operon and is often encoded by an adjacent hisR gene, suggesting it`s likely acquisition via horizontal gene transfer. The predicted DNA binding motifs of HisR in various taxonomic groups have a structure of 20-bp palindromes with a common consensus sequence YACTTTANYNNRNTAAAGTR, where Y is C or T, S is C or G, and W is A or T, and has some variations in several conserved positions in certain taxa.

Acknowledgments

This research was supported by the Russian Science Foundation (grant #19-14-00305). Natalia V. Sernova is grateful to Dr. Victor A. Kostyuchenko for the priceless help with Linux-script that has made a site-search procedure much faster.

References

1. Alifano, P., Fani, R., Lio, P., Lazcano, A., Bazzicalupo, M., Carlomagno, M.S., Bruni, C.B.: Histidine biosynthetic pathway and genes: structure, regulation, and evolution. Microbiological reviews 60, 44-69 (1996)
2. Winkler, M.E., Ramos-Montanez, S.: Biosynthesis of Histidine. EcoSal Plus 2009, (2009)
3. Vitreschak AG, Lyubetskaya EV, Shirshin MA, Gelfand MS, Lyubetsky VA. Attenuation regulation of amino acid biosynthetic operons in proteobacteria: comparative genomics analysis. FEMS Microbiol Lett. 234(2), 357-70 (2004)
4. Leyn, S.A., Kazanov, M.D., Sernova, N.V., Ermakova, E.O., Novichkov, P.S., Rodionov, D.A. Genomic reconstruction of the transcriptional regulatory network in *Bacillus subtilis*. J Bacteriol. 195(11), 2463-73 (2013)
5. Ravcheev, D.A., Best, A.A., Tintle, N., Dejongh, M., Osterman, A.L., Novichkov, P.S., Rodionov, D.A. Inference of the transcriptional regulatory network in *Staphylococcus aureus* by integration of experimental and genomics-based evidence. J Bacteriol. 193(13), 3228-40 (2011)
6. Vitreschak, A.G., Mironov, A.A., Lyubetsky, V.A., Gelfand, M.S. Comparative genomic analysis of T-box regulatory systems in bacteria. RNA. 14(4), 717-35 (2008)
7. Sun, E.I., Leyn, S.A., Kazanov, M.D., Saier, M.H. Jr, Novichkov, P.S., Rodionov, D.A. Comparative genomics of metabolic capacities of regulons controlled by cis-regulatory RNA motifs in bacteria. BMC Genomics. 14, 597 (2013)

8. Overbeek, R., Begley, T., Butler, R.M., Choudhuri, J.V., Chuang, H.Y., Cohoon, M., de Crecy-Lagard, V., Diaz, N., Disz, T., Edwards, R., Fonstein, M., Frank, E.D., Gerdes, S., Glass, E.M., Goesmann, A., Hanson, A., Iwata-Reuyl, D., Jensen, R., Jamshidi, N., Krause, L., Kubal, M., Larsen, N., Linke, B., McHardy, A.C., Meyer, F., Neuweger, H., Olsen, G., Olson, R., Osterman, A., Portnoy, V., Pusch, G.D., Rodionov, D.A., Ruckert, C., Steiner, J., Stevens, R., Thiele, I., Vassieva, O., Ye, Y., Zagnitko, O., Vonstein, V.: The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic acids research 33, 5691-5702 (2005)
9. Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E.L., Tate, J., Punta, M.: Pfam: the protein families database. Nucleic acids research 42, D222-230 (2014)
10. UniProt, C.: UniProt: a hub for protein information. Nucleic acids research 43, D204-212 (2015)
11. Saier, M.H. Jr, Reddy, V.S., Tsu, B.V., Ahmed, M.S., Li, C., Moreno-Hagelsieb, G. The Transporter Classification Database (TCDB): recent advances. Nucleic Acids Res. 44(D1), D372-9 (2016)
12. Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O.: New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic biology 59, 307-321 (2010)
13. Novichkov, P.S., Rodionov, D.A., Stavrovskaya, E.D., Novichkova, E.S., Kazakov, A.E., Gelfand, M.S., Arkin, A.P., Mironov, A.A., Dubchak, I.: RegPredict: an integrated system for regulon inference in prokaryotes by comparative genomics approach. Nucleic acids research 38, W299-307 (2010)
14. Rodionov, D.A. Comparative genomic reconstruction of transcriptional regulatory networks in bacteria. Chemical reviews 107, 3467-3497 (2007)
15. Katoh, K., Standley, D.M.: MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular biology and evolution 30, 772-780 (2013)
16. Mukherjee, S., Sengupta, S. Riboswitch Scanner: an efficient pHMM-based web-server to detect riboswitches in genomic sequences. Bioinformatics. 32(5), 776-8 (2016)
17. Kalvari, I., Nawrocki, E.P., Argasinska, J., Quinones-Olvera, N., Finn, R.D., Bateman, A., Petrov, A.I. Non-Coding RNA Analysis Using the Rfam Database. Curr Protoc Bioinformatics. 62(1), e51 (2018)