EVOLUTIONAL AND FUNCTIONAL ANALYSIS OF T-BOX REGULON IN BACTERIA: IDENTIFICATION OF NEW GENES INVOLVED IN AMINO ACID METABOLISM

***Vitreschak A.G.*[*1], *Lyubetsky V.A.*[1], *Gelfand M.S.*[1]**
[1] Institute for Information Transmission Problems, RAS, Moscow, 127994, Russia,
[*] Corresponding author: l_veter@mail.ru

Key words: genome analysis, amino acid biosynthesis and transport, T-box, bacteria

SUMMARY

Motivation: T-box antitermiantion is the most distributed mechanism of regulation of various amino acids in Gram-positive bacteria. Identification of the T-box regulon and a metabolic analysis of amino acid biosynthesis and transport is one of problems of comparative genetics, genomics and molecular biology.

Results: Search for T-box elements and analysis of operon structures identified a large number of new candidate T-box regulated genes, mostly transporters, in Gram-positive bacteria. We assign amino acid specificity for a large number of candidate transporters as well as for other new amino acid related genes.

Availability: The program is available by request to the author.

INTRODUCTION

Computer comparative analysis is a powerful method of prediction of the RNA secondary structure. It has been used for prediction of both regulatory and structural RNAs. A somewhat different approach is to predict gene regulation by analysis of RNA patterns. We have used it to analyze the T-box regulatory elements in Gram positive bacteria. It is experimentally known a number of T-box elements in some Gram positive bacteria: *Bacilllus subtilis*, *Bacillus stearothermophilus*, *Lactococcus lactis* and *Staphylococcus aureus* and some others (Grundy *et al*., 1994; and others). Genes is known to be regulated by T-boxes encode in most cases aminoacyl-tRNA synthetases, amino acid biosynthetic operons and some amino acid transporters. The T-box regulatory element consists of the alternative RNA secondary structures (the terminator and antiterminator conformations) and a number of conservative sequences boxes. The uncharged amino acid-tRNA is the inducer of transcription. At low concentration of regulatory amino acid in medium it binds to the RNA structure (interacts with T-box and anti-anticodon site) and promotes formation of the antiterminator. In contrast, at high concentration of regulatory amino acid a terminator conformation forms that leads to premature termination of transcription.

DATA AND METHODS

Using the set of known T-box sites, we constructed the pattern of the T-box RNA element and scanned available genomic sequences using the RNA-PATTERN program

(Vitreschak *et al*., 2001) and another program, developed for these purposes (Leontiev, Lyubetsky, 2006). The input RNA pattern described the RNA secondary structure and the sequence consensus motifs. The RNA secondary structure was described as a set of the following parameters: the number of helices, the length of each helix, the loop lengths and the description of the topology of helix pairs.

RESULTS AND DISCUSSION

We found about 800 T-boxes in 90 bacterial genomes. T-boxes are widely distributed in Gram-positive bacteria (Firmicutes, Actinobacteria). Moreover, several T-boxes were found in some Gram-negative bacteria (δ-proteobacteria) and other groups (Dienococcales\Thermales, Chloroflexi, Dictyoglomi).

Comparison of sets of T-box-regulated genes in analysed genomes shows, that most genes is constituted by aminoacyl-tRNA synthetase genes. Two other groups of T-box regulated genes consist of amino acid biosynthetic genes and genes with unknown function. Distribution of T-boxes involved in regulation of aminoacyl-tRNA synthetase and amino acid biosynthetic genes by T-box antitermination is shown in Table 1.

Aminoacyl tRNA synthetase genes *ileS*, *valS*, *leuS*, *serS*, *thrS*, *pheST*, *alaS*, *asp*(*asn*)*S*, *glyS*(*QS*) are regulated by T-box antitermination in most Firmicutes and some other phylogenetic groups, whereas *metS*, *proS*, *CyS*, *hisS*, *argS*, *lysS* are regulated only in distinct groups/bacteria.

T-box antitermination mechanism is also involved in regulation of various amino acid biosynthetic genes. *trp* and *ilv(leu)* operons are found to be regulated in most Firmicutes as well as in some other groups. Other amino acid biosynthetic genes are regulated only in distinct groups/bacteria (Table 1). The conservation of the T-box antitermination in distinct groups can be explained by a variability of regulatory mechanisms. In particular, the methionine metabolism in Gram-positive bacteria was known to be controlled by five different mechanisms: S-box, T-box, metK-box regulation (acting on the level of premature termination of transcription/inhibition of translation initiation) and two other mechanisms acting on the DNA level (Met-box and MetJ-box) (Rodionov *et al*., 2004). In another case, regulation of genes of the aromatic amino acid biosynthesis pathway in Gram-positive bacteria is shown to be quite labile and involves at least four regulatory systems, two at the RNA level involving competition of alternative RNA secondary structures for transcription and/or translation regulation and two at the DNA level (Panina *et al*., 2003).

Positional analysis of T-boxes led to the identification of a large number of new candidate amino acid transporters (Table 2).

We predicted the amino acid specificity of possible transporters analyzing the T-box regulatory "specifier codon" (a T-box regulatory site involved in the interaction with the anti-codon site of the uncharged tRNA). The regulatory codon of the T-box RNA element is known to be located in the fixed internal loop of the specifier hairpin. We verified the amino acid specificity of all predicted T-boxes was by sequence and structural alignment (multAl, Mironov, unpublished) and construction of phylogenetic trees (In most cases, T-boxes with the same specificity located in the same branch of the T-box phylogenetic tree).

The predicted tyrosine specific transporter *yheL* (Na+/H+ antiporter) is found to be regulated by the (TYR)T-box antitermination in some Bacillales and Lactobacillales. A phylogenetic analysis showed that YheL form a separate branch on the NhaC superfamily phylogenetic tree. This family also includes lysine transporters LysW, methionine transporters MetT and malate/lactate antiporter MleN.

Table 1. Regulation of aminoacyl-tRNA synthetases and amino acid biosynthetic operons in Gram-positive bacteria

Amynoacyl-tRNA synthetases	
Aromatic a/a TRP, PHE, TYR	Most FIRMICUTES, *Atopobium minutum*
Branched chain a/a ILE, LEU,VAL	Most FIRMICUTES, Actinobacteria(ileS), Dienococcales\ Thermales(ileS, valS), Chloroflexi(ileS), *Thermomicrobium roseum*(leuS)
methionine	Bacillales, Clostridiales, *Thermoanaerobacter tengcongensis*
proline	Some Bacillales, Clostridiales,
cysteine	Bacillales, some Lactobacillales, Clostridiales, Thermoanaerobacteriales
histidine	Bacillales, Lactobacillales(exept streptococcus spp.), some Clostridiales, *Thermoanaerobacter tengcongensis*
arginine	Bacillales, Lactobacillales (exept streptococcus spp.), Clostridiales,
threonine	Bacillales, Lactobacillales, Clostridiales, Dictyoglomi, Thermomicrobium roseum
serine	Most FIRMICUTES
alanine	Bacillales, Lactobacillales, Clostridiales
ASP, ASN	Most FIRMICUTES (exept streptococcus spp., Mycoplasmatales, Entomoplasmatales)
glycine	Most FIRMICUTES, Dienococcales\ Thermales
lysine	*Bacillus cereus*, *Clostridium thermocellum*
Amino acid biosynthetic genes	
Aromatic a/a TRP, PHE, TYR	Most FIRMICUTES, Chloroflexi and Dictyoglomi (trp operon), some FIRMICUTES (aro genes, pheA, pah)
Branched chain a/a ILE, LEU,VAL	Bacillales, Clostridiales, *Syntrophomonas wolfei*, δ-proteobacteria(leu), Dictyoglomi, *Thermomicrobium roseum*
methionine	Lactobacillales (exept streptococcus spp.), *Desulfotomaculum reducens*
proline	Bacillales, Desulfitobacterium hafniense, *Desulfotomaculum reducens*
cysteine	Bacillales, Enterococcus faecalis, *Clostridium acetobutylicum*, Dictyoglomi
histidine	some Lactobacillales
arginine	*Clostridium difficile*
threonine	Bacillus cereus, *Clostridium difficile*
serine	some FIRMICUTES
alanine	–
ASP, ASN	some FIRMICUTES
glutamine	*Clostridium perfringes*
glycine	–
lysine	–

In addition to two known tryptophan transporters, *yhaG* and *ycbK*, two new tryptophan transport systems were identified: *trpXYZ* (Peptococcaceae, *Streptococcus spp*., *Paenibacillus larvae*) and *yocR*(*yhdH*)(*Bacillus cereus*).

New large family of amino acid ABC transporters was characterized. In addition to previously described methionine ABC transporter *yusCBA* (Zhang *et al*., 2003) we found five new amino acid ABC transporters from this ABC transporter superfamily: *yqiXYZ*(*ARG*), *hisXYZ*(*HIS*), *yckKJI*(*CYS/MET*), *aspQHMP*(*ASP*), *ytmKLM*(*MET*).

The specificity of various possible amino acid permeases was predicted: *yvbW*(*LEU*), *ykbA*(*THR*), *lysX*(LYS), *RDF02391*(*ARG*).

Genes encoding transporters from branched-chain amino acid transporter family was found to be regulated by three amino acids: ILE (some Bacillales, Lactobacillales and Clostridiales), VAL(some Lactobacillales), THR (*Bacillus cereus, Clostridium tetani*).

Analysis of the methionine-specific T-box regulatory signals allowed us to identify hypothetical oligopeptide ABC transport system in Gram-positive bacteria, *opp*, which is possibly involved in the uptake of some methionine precursors or oligopeptides.

Table 2. Regulation of amino acid transporters by T-box antitermiantion in Gram-positive bacteria

Gene	Sp.	Predicted function	Bacteria
ycbK	TRP	tryptophan-specific permease	*Bacillus subtilis, Bacillus licheniformis*
yhaG	TRP	tryptophan-specific permease	Clostridiales
yvbW	LEU	leucine-specific permease	*Bacillus subtilis, Bacillus licheniformis*
ykbA	THR	threonine-specific permease	*Bacillus subtilis*
ybgF/aapA	?	?	*Lactobacillus reuteri*
yheL	TYR	Tyrosine transporter (Na+/H+ antiporter)	some Bacillales and Lactobacillales
lysX	LYS	lysine transporter	some Bacillales
	ILE	Branched-chain amino acid transporter family: ILE-specific	some Bacillales, Lactobacillales andClostridiales
brnQ_braB	THR	Branched-chain amino acid transporter family: THR-specific	*Bacillus cereus, Clostridium tetani*
	VAL	Branched-chain amino acid transporter family: VAL-specific	some Lactobacillales
yusCBA	MET	methionine ABC transporter	Lactobacillales, *Enterococcus faecalis*
yqiXYZ	ARG	arginine ABC transporter	*Clostridium difficile*
hisXYZ	HIS	histidine ABC transporter	Lactobacillales, *Clostridium difficile*, *Listeria monocytogenes*, *E. faecalis*
yckKJI	CYS	cysteine ABC transporter	*Clostridium acetobutylicum*
	MET	methionine ABC transporter	some Lactobacillales
aspQHMP	ASP	ASP(ASN) ABC transporter	*Lactobacillus johnsonii*
ytmKLM	MET	methionine ABC transporter	*Leuconostoc mesenteroides*
	TRP	TRP-*specific sodium dependent transporter*	*Bacillus cereus*
	PHE	PHE-*specific sodium dependent transporter*	*Bacillus cereus*
	LEU	LEU-*specific sodium dependent transporter*	*Bacillus cereus*
	?	*sodium dependent transporter*	*Clostridium tetani*
mtsABC opp	MET	uptake of unknown methionine precursors, possibly oligopeptides	some Lactobacillales
trpXYZ	TRP	tryptophan ABC transporter	Peptococcaceae, *Streptococcus spp.*, *Paenibacillus larvae*
RDF02391	ARG	arginine permease	*Clostridium difficile*
ABC-like	?	?	*Desulfotomaculum reducens*
CBX	?	?	*Clostridium botulinum*
gltT like	?	?	*some Clostridium spp.*

New possible amino acid transporters are in bold. Predicted specificity of an amino acid transporter is shown in second column.

AKNOWLEGEMENTS

This study was partially supported by grants CDF RBO-1268 from the Ludwig Institute for Cancer Research and 55000309 from the Howard Hughes Medical Institute.

REFERENCES

Grundy F.J., Rollins S.M., Henkin T.M. (1994) Interaction between the acceptor end of tRNA and the T box stimulates antitermination in the *Bacillus subtilis* tyrS gene: a new role for the discriminator base. *J. Bacteriol.*, **176**, 4518–4526.

Leontiev L.A., Lyubetsky V.A. (2006) Massive search of conserved regulatory structures containing T-boxes: results of calculation. *Information processes*, **6**, 20–23.

Panina E.M., Vitreschak A.G., Mironov A.A., Gelfand M.S. (2003) Regulation of biosynthesis and transport of aromatic amino acids in low-GC Gram-positive bacteria. *FEMS Microbiol. Lett.*, **28**, 211–220.

Rodionov D.A., Vitreschak A.G., Mironov A.A., Gelfand M.S. (2004) Comparative genomics of the methionine metabolism in Gram-positive bacteria: a variety of regulatory systems. *Nucl. Acids Res.*, **32**, 3340–3353.

Vitreshchak A.A., Mironov A.A., Gelfand M.S. (2001) Computer prediction of RNA secondary structure. The RNApattern program: searching for RNA secondary structure by the pattern rule. In *Proceedings of the Third International Conference "ComplexSystems: Control and modeling problems*". Russia, Samara, pp. 623–255.

Zhang Z., Feige J.N., Chang A.B., Anderson I.J., Brodianski V.M., Vitreschak A.G., Gelfand M.S., Saier M.H. Jr. (2003) A transporter of *Escherichia coli* specific for L- and D-methionine is the prototype for a new family within the ABC superfamily. *Archives of Microbiology*, **180**, 88–100.