

Conservation of the binding site for the arginine repressor in all bacterial lineages

ABSTRACT

We found that the ArgR/AhrC recognition signal is conserved in all genomes of orthologous transcription factors of this family, with all but the *M. tuberculosis* genome having candidate arginine regulons (which are associated with *E. coli*'s Art system).

INTRODUCTION

The primary function of the repressor is to prevent the formation of an insoluble binding site for the arginine repressor. The binding site of arginine repressor is located in the cytoplasmic membrane (CWM). This CWM is an extremely important site of protein-protein interactions. Because of its importance, the CWM is considered to be the main site of interaction among bacterial proteins [1,2].

The CWM is also the site of the interaction between arginine and phosphatidylinositol-3-phosphate (PI3K). PI3K is a major regulator of protein synthesis in all bacterial lineages [3]. PI3K inhibition leads to the inhibition of protein synthesis in many bacterial lineages, including streptococci. Bacterial and archaeal transcriptional regulators commonly form large protein families that consist of numerous paralogs, such as LacI/GntR, AraC, and DeoR. However, only three or two representatives from a diverse range of bacteria are easily detectable among the orthologous transcription factors found in clusters (Table 1). Comparisons of the coevolution of these regulator species and their binding sites in DNA may reveal general patterns in the evolution of regulons. In Gram-negative bacteria, LexA and DinR have distinct signals, while their respective DNA-binding domains are separated (Table 1). In contrast, the arginine regulon, which is controlled by the CIRCE element and not the SOS or heat-shock regulon, represents an evolutionary strategy. Using the *E. coli* regulon, which was previously well-characterized, we used comparative genomic analysis to predict the gene composition of the arginine regulon of *Haemophilus influenzae* using regulatory signals. We now examine the conservation of this ARG box in all bacteria that encode an ortholog of these ArgR repressors.

CONCLUSION

Remarkable conclusions The differences in the composition of ARG regulons among bacteria are mainly due to differences that occur in different pathways of arginine degradation and species-specific paralogs. This leads to the question of where the additional ARB boxes come from, as some sites may not be convergent and others may be located upstream of an ArgR-binding signal (e.g. On the contrary, CIRCE elements seem to be direct descendants of the ancient regulon found in the common ancestor of Bacteria, as little variation exists in what is composed of this regulon and only a few extra sites in some genomes appear to have been product of duplication; most other DNA-binding domains of transcriptional regulators (such as LexA) seem now to undergo considerable changes with their signal and regulation but they no longer reflect all genetic information that must be expressed. The use of single and cooperative sites in the arginine regulon strategy is noteworthy. Although this concept may not be directly addressed in this study, it seems that the need for a sharper response to ArgR's stimulus (examples of which are usually found

in *E. coli*) than to SOS may have led to the formation of cooperative binding sites, which cannot be easily summarized using systematic conclusions due to weaker second sites within the cooperative cassettes. Among the other transcription factor families, there are several (biotin operon repressor, COG1654; putative stress-responsive transcriptional regulator PcpC, in COG1983; Bvg accessory factor homologs, CDM1521) with one representative each genome that we would like to compare, but they do not have enough experimentally determined binding sites and are not as common among the three regulators described above. Our approach, combined with positional analysis of recently discovered co-localized enzymes and regulator genes, is convincing enough for testing whether complex molecules still remains