# In silico discovery of the dormancy regulons in a number of Actinobacteria genomes

Anna Gerasimova[1,2*], Inna Dubchak[1,2], Adam Arkin[1,2], Mikhail Gelfand[3]

[1]Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA; [2]Energy Biosciences Institute , Berkeley, CA 94720, USA; [3]Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow 127994, Russia

* AGerasimova@lbl.gov

## Introduction

*Mycobacterium tuberculosis* is a dangerous Actinobacteria infecting nearly one third of the human population. It becomes dormant and phenotypically drug resistant in response to stresses. An important feature of the *M. tuberculosis* pathogenesis is the prevalence of latent infection without disease, making understanding of the mechanisms used by the bacteria to exist in this state and to switch to metabolically active infectious form a vital problem to consider.

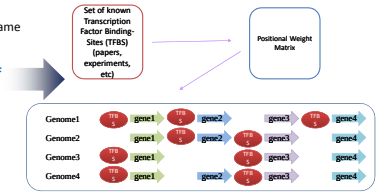
*M. tuberculosis* dormancy is regulated by the three-component regulatory system of two kinases (DosT and DevS) and transcriprional regulator (DevR). DevR activates transcription of a set of genes, which allow the bacteria to survive long periods of anaerobiosis, and may be important for long-term survival within the host during latent infection. The DevR-regulon is studied experimentally in *M. tuberculosis* and few other phylogenetically close *Mycobacteria* spp. As many other two-component systems, the *devRS* operon is autoregulated. However, the mechanism of the dormancy is not completely clear even for these bacteria and there is no data describing the dormancy regulons in other species.

## Data & Methods

**Regulon** - set of genes in a particular genome that are directly regulated by the same transcription factor

**Our approach: increasing sensitivity of site detection by comparing multiple species**

**Regulog** - group of regulons operated by the orthologous transcription factors in closely-related genomes




## Results

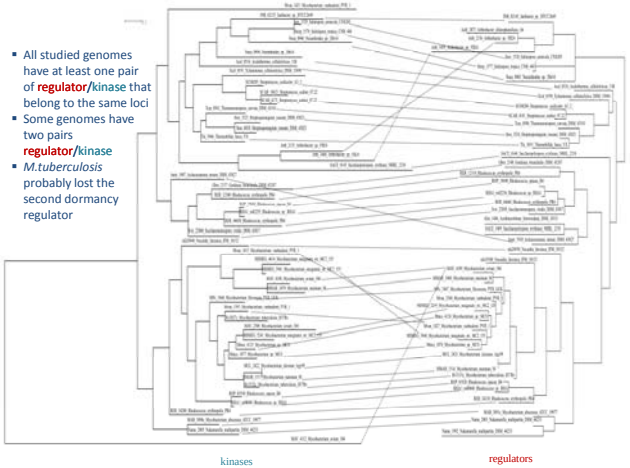
### We determined *Actinobacteria* spp that may be dormant

Species tree by MicrobesOnline

### We detected duplication of the regulatory systems in *M.marinum, M.smegmatis, M.vanbaalenii, Rhodococcus* spp, *S. roseum* and *Arthobacter* sp FB24

### Co-localisation of regulators and kinases in *Actinobacteria*

- All studied genomes have at least one pair of regulator/kinase that belong to the same loci
- Some genomes have two pairs regulator/kinase
- *M.tuberculosis* probably lost the second dormancy regulator




Kinases and regulators that belong to the same *loci* are bound by a dotted line

### For all studied genomes we have constructed the DevR-binding motif

### With the PWM for the DevR-motif we scanned the genomes and described dormancy regulog

### DevR-regulation of regulators (devR), kinases (devS) and genes in the adjacent loci

The identified DevR-regulons contain many paralogs. For example, the 22 Kbp *devR* loci may contain up to five DevR-regulated universal stress proteins (USPs). We identified additional DevR-regulated USPs, nitrogenases and diacylglycerol-acyltransferases (DGATs). On the other hand, many USPs, nitrogenases, and DGATs do not seem to be DevR-regulated. To study the evolution of DevR regulation in these large gene families, we constructed pylogenetic trees and analyzed the distribution of regulated genes. (data not shown)

## Conclusions

We successfully applied comparative genomic analysis methods to characterize the DevR-regulons in thirty five Actinobacteria genomes.

We also managed to predict new regulon members in *M. bovis*, *M.smegmatis* and *M.avium*. We defined the core of the dormancy regulons in most species analyzed and detected regulon members that are specific for small groups of genomes only. In most studied species the proteins involved in anaerobic respiration, such as: nitroreductases, cytochrome and NADH-oxidoreductases, ferredoxin and nitrate/nitrite transporters turn out to contain DevR-binding motifs. The DevR-regulation of heat shock proteins is also conserved in most genomes researched.

Moreover, all studied Actinobacteria contain universal stress proteins that are under dormancy regulation. Furthermore, we have carried out functional evolutionary analysis of the regulators and large families of regulated proteins, and were able to detect duplication of the regulatory system in *M.marinum, M.smegmatis*, *M.vanbaalenii*, Rhodococcus spp, *S. roseum* and *Arthobacter* sp FB24. The autoregulation was conserved in all studied species except *S. erythraea* and *R. opacus*. Hence, the most conservative members of dormancy regulog are USPs and the DevR/DevS regulators.

## Acknowledgments

We are grateful for software for genome analysis to **Pavel Novichkov, Elena Stavrovskaya** and **Andrei Mironov** (**RegPredict** team) and **Morgan Price**, **Paramvir Dehal** and **Keith Keller** (**MicrobesOnline** team) and for useful discussions to **Alexey Kazakov** and **Gregory Dolganov**.

This work was supported by the Director, Office of Science, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231

## LEGAL DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California.