Phylogenomic analysis of natural products biosynthetic gene clusters allows discovery of arseno-organic metabolites in model streptomycetes

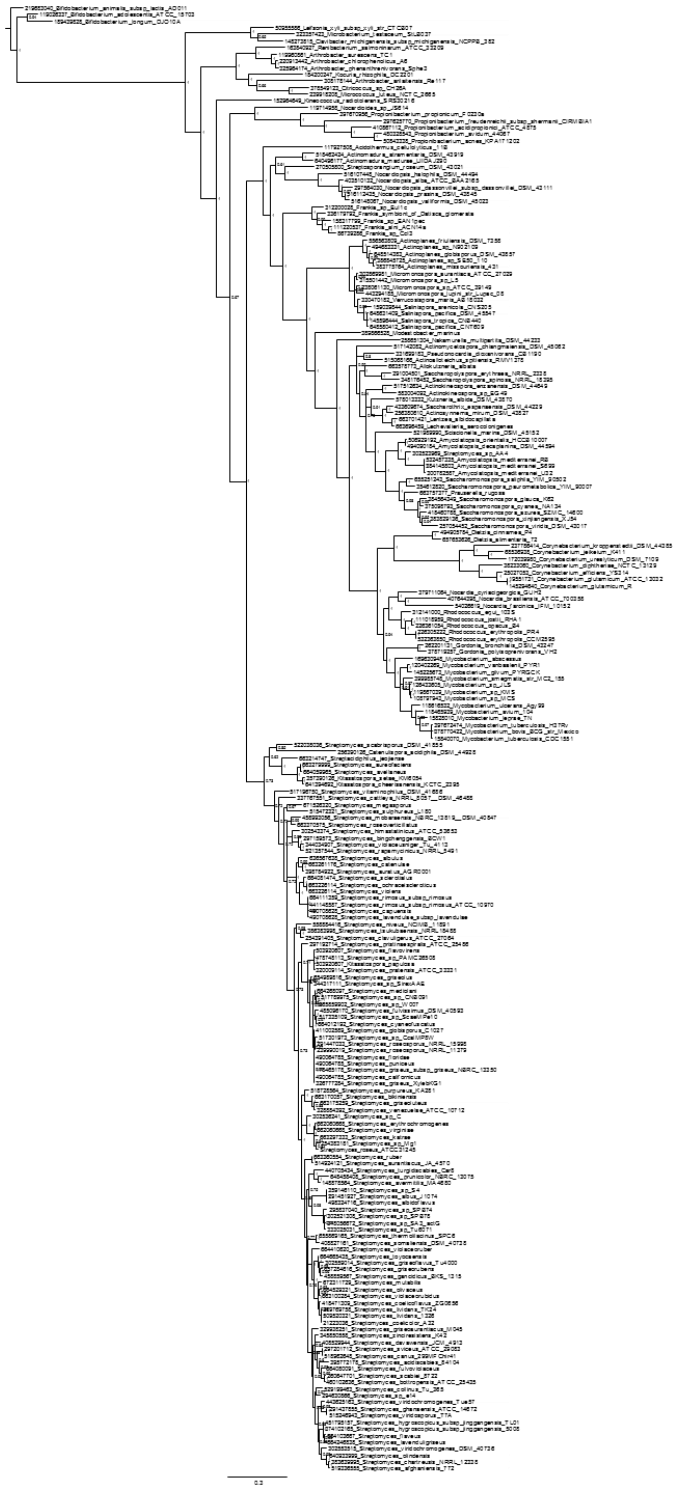
Pablo Cruz-Morales[1,*], Johannes Florian Kopp[2], Christian E. Martínez-Guerrero[1], Luis Yáñez-Guerra[1], Nelly Selem-Mojica[1], Hilda E. Ramos-Aboites[1], Jörg Feldmann[2], & Francisco Barona-Gómez[1,*]

1 Evolution of Metabolic Diversity Laboratory, Langebio, Cinvestav-IPN. Irapuato, Guanajuato, México.

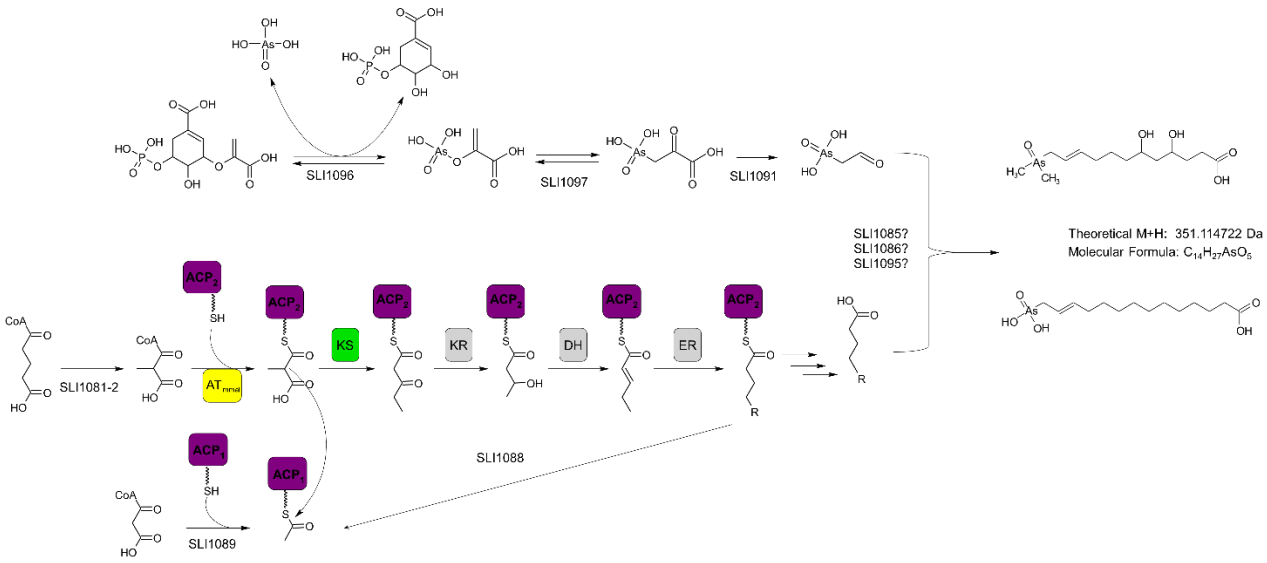
2 Trace Element Speciation Laboratory (TESLA), College of Physical Sciences. Aberdeen, Scotland, UK.

*Authors for correspondence:

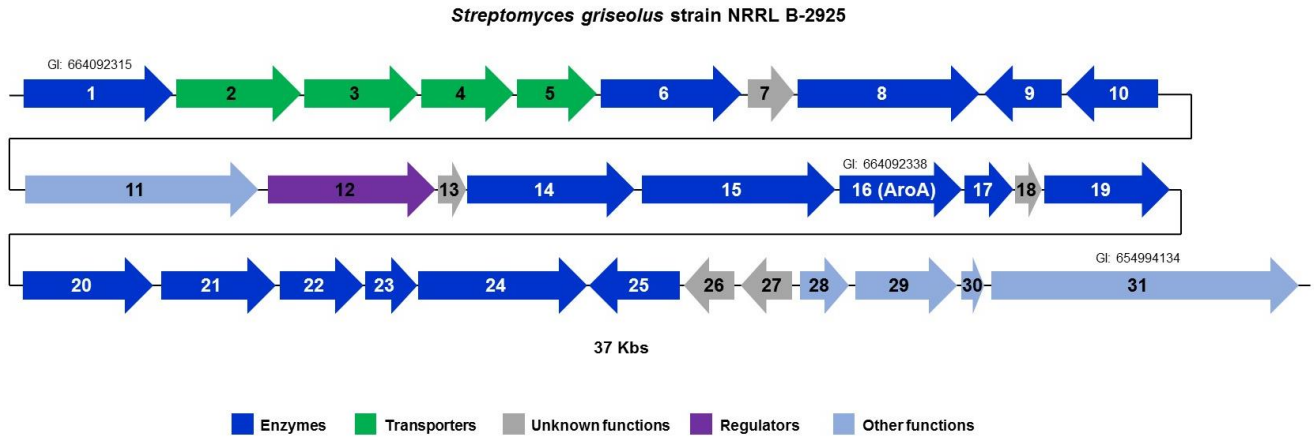
Francisco Barona-Gómez (fbarona@langebio.cinvestav.mx)
Pablo Cruz-Morales (pcruz@langebio.cinvestav.mx)
Evolution of Metabolic Diversity Laboratory,
Langebio, Cinvestav-IPN,
Irapuato, Guanajuato, México



Supplementary figure 1. Species tree representing the taxonomic relationships of the strain in the genome database. The phylogenetic tree was constructed the RNA polymerase Beta subunit (RPOB) amino acid sequence.



Supplementary figure 2. Postulated pathway for arseno-organic NP biosynthesis in *S. coelicolor* and *S. lividans.* The reactions proposed for SLI_1096, SLI_1097 and SLI_1091 are responsible for the biosynthesis of the As-C bond at the early stages of the biosynthetic pathway. The biosynthetic logic proposed for SLI_1088-9 is related to the synthesis of an acyl chain that is proposed to be linked to the As-C containing intermediary by other enzymes in the BGC. At the left, structural predictions of potential products for the pathway based on high resolution MS data are shown. This pathway and further studies on the water-soluble As-species present in the samples (data not shown) suggest a non-methylated As-moiety as shown in the last structure, which has not been described in literature yet.



Supplementary figure 3. Genetic structure of a putative no-validated BGC conserved in *Actinobacteria*. This BGC was predicted after identification of a recruited AroA homolog which was not identified by ClusterFinder or antiSMASH. Detailed annotation is available as supplementary table S8.