Genome mining: an approach to finding bioactive secondary metabolites in *Streptomyces* sp. strain K155.

Guzmán-Trampe S, Ikeda H[†] and Sánchez S. Departamento de Biología Molecular y Biotecnología. Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, México D.F. 04510.
[†]Kitasato Inst. for Life Sci., Kitasato Univ. Kanagawa 252-0373, Japan.
E-mail: no_encontre@hotmail.com.

Key words: secondary metabolites, genome mining, Streptomyces

Introduction. *Streptomyces* are high-GC content Gram-positive bacteria that during their life cycle form filamentous mycelia, aerial hyphae and spores. Species on this genus produce a variety of secondary metabolites and holds the production of two thirds of the pharmaceutical compounds produced by bacteria (1). Genome driven experiments have shown an unexpected number of genes involved in biosynthesis of bioactive compounds that remained "cryptic" until now. Genome mining of *Streptomyces* has revealed the presence of more than 20 gene clusters involved in secondary metabolites production like PKS, NRPS, terpenoid, aminoglycosides and shikimate-derived metabolites (2). Previous work has shown the ability of *Streptomyces sp.* strain K155 to produce antibiotic compounds (3). Since this strain seems to be a new species (data not shown), we are interested in addressing a genomic approach to evaluate its potential for secondary metabolites production.

Methods. DNA extraction and purification was performed. Sequencing and *de novo* assembly was conducted by BaseClear. Annotation was made with MetaGeneAnnotator. BLAST and Pfam was applied to confirm protein-coding genes. To explore and define putative gene clusters for secondary metabolites in this strain, protein sequences from 37 and 38 biosynthetic gene clusters from *S. avermitilis* and *S. griseus*, respectively (2) were used. With the same approach, the presence of the main primary metabolism genes necessary for secondary metabolism in K155 genome, were examined and compared with *S. coelicolor*, *S. griseus* and *S. avermitilis* genes.

Results. The genome contained at least 7,373,781 bases, averaged a GC content of 71.53% and encoded at least 6,966 potential proteins. Although the sequence of *Streptomyces* sp. K155 has not been completely annotated, BLAST and Pfam data revealed thirteen gene clusters related with biosynthesis of ectoine, nocardamine, terpenes, a lantibiotic, siderophores and type I and II PKS's (Table 1). This means that 1.13% of the *Streptomyces* sp. K155 genome is devoted to synthesis of secondary metabolites. Until now, biosynthetic gene clusters for NRPS have not been found.

Table 1. Secondary metabolite gene clusters in *Streptomyces* sp. K155.

No.	Cluster location	Predicted product
1	K155IL_00432-K155IL_00437	gene cluster for nocardamine biosynthesis
2	K155IL_00513- K155IL_00516	gene cluster for ectoine synthase
3	K155IL 01196- K155IL_01200	gene cluster for diterpene production
4	K155IL 02188-K155IL_02203	gene cluster for type II PKS synthesis
5	K155IL 02229-K155IL_02251	generation of modified dioxyaminosugar
6	K155IL 02389-K155IL_02396	gene cluster for spore pigment biosynthesis
7	K155IL_03437	4-hydroxyphenylpyruvate dioxygenase
8	K155IL 03584-K155IL_03587	gene cluster for siderophore synthesis
9	K155IL 04080-K155IL_04084	gene cluster for squalene/hopanoid synthesis
10	K155IL_04686- K155IL_04687	gene cluster for albaflavenone biosynthesis
11	K155IL 05334-K155IL_05336	lantibiotic operon
12	K155IL_05994	putative type-I PKS, truncated
13	K155IL_6022	germacradienol/geosmin synthase

The number of predicted genes encoding transfer RNAs was at least 56. We also found genes encoding transposase in many regions. As secondary metabolism is regulated by the availability of precursors synthesized by primary metabolism, we explored for the presence and copy numbers of the main genes involved carbon, and amino acid pathways. We found 41 genes involved in primary metabolism: *glk*-2, *pgi*-3, *pfkA*-3, *fba*-1, *gap*-2, *pgk*-1, *gpmA*-2, *eno*-1, *pykA*-2, *aceB*-1, *aceE*-2, *aceF*-1, *poxB*-2, *citA*-2, *acnA*-1, *icdA*-1, *sucA*-1, *sucB*-1, *ipdA*-2, *ipdB*-1, *korA*-1, *korB*-1, *sdhA*-3, *sdhB*-3, *sdhC*-2, *sdhD*-1, *sucC*-3, *sucD*-2, *fumB*-1, *fumC*-1, *mdh*-1, *malS*-2, *pckA*-1, *tkt*-2, *tal*-1, *zwf*-2, *pgl*-1, *gnd*-3, *rpiB*-2 and *rpe*-1 (copy number of each gene is presented). Although the specific function of these genes is unknown, specific deletion leads to over or under expression of the secondary metabolites.

Conclusions. The results of *Streptomyces* sp. K155 genome are similar (size and GC content) to those reported for other *Streptomyces* strains. Until now thirteen biosynthetic gene clusters have been revealed but their metabolic products remain to be identified experimentally.

References.

1. Demain A. y Sánchez S. 2009. *J Antibiot*. 62: 5-16.
2. Nett M, Ikeda H and Moore S. 2009. Nat. Prod. Rep. 26: 1362-1384.
3. Ávalos M. 2010. Undergraduate thesis, Facultad de Química, UNAM.