

Searching For Specialized Metabolites Gene Clusters By Mining Genomic Islands: A New Comparative Genomics-Based Approach

Drago Haas¹, Matthieu Barba¹, Cláudia M. Vincente², Géraldine Le Goff³, Jamal Ouazzani³, Bertrand Aigle², Pierre Leblond², Olivier Lespinet¹, Jean-Luc Pernodet¹,
Sylvie Lautru^{1*}

¹ Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ.Paris-Sud, University Paris-Saclay, France.

² Dynamique des génomes et adaptation microbienne (DynAMic), UMR 1128, INRA, Université de Lorraine, France.

³ Institut de Chimie des Substances Naturelles (ICSN), CNRS, Gif-sur-Yvette, France

Microbial specialized metabolism biosynthetic gene clusters (SMBGCs) are a rich source of natural products of medical interest. With the multiplication of genomic data now available, bioinformatic tools for automatic SMBGCs detection have been developed in the last few years. These tools (e.g. antiSMASH [1]) are able to identify typical specialized metabolism genes by searching for homology regions with well characterized genes and have led to the discovery of several new bioactive compounds [2]. However, the discovery of atypical SMBGCs encoding under-characterized enzymes remains difficult.

SMBGCs are known to be often acquired by horizontal gene transfer leading to the assumption that genomic islands (GIs) are enriched in secondary metabolism genes. Studies have already shown good correlation between the SMBGC and GI profiles, e.g. in *Salinispora* genomes [3]. Here we present the Synteruptor tool. Synteruptor is a user-friendly software developed for the identification of GIs in closely related organisms. It performs a synteny comparison of two genomes allowing the detection of synteny breaks. To illustrate the use of Synteruptor, a comparison of *Streptomyces ambofaciens* and *Streptomyces coelicolor* was performed. This allowed the identification of known SMBGCs as well as that of an atypical one involved in the synthesis of several metabolites, among which sphydrofurane, previously isolated the culture media of several *Streptomyces*.

To our knowledge, Synteruptor is the first user-friendly SMBGC search tool not based on the identification of genes by homology. We have shown that it is able to detect typical SMBGCs as well as atypical ones, as illustrated by the discovery of the sphydrofuran biosynthetic gene cluster in *Streptomyces ambofaciens*. Synteruptor is complementary to the currently used tools such as AntiSMASH and can be use in synergy with them.

Keywords: Genomic islands mining;

References

[1] 1. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH. (2015). antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* 43(W1):W237-243.

[3] Corre C, Challis GL (2009). New natural product biosynthetic chemistry discovered by genome mining. *Nat. Prod. Rep.* 26:977-986

[2] Ziemert N, Lechner A, Wietz M, Millán-Aguiñaga N, Chavarria KL, Jensen PR. (2014). Diversity and evolution of secondary metabolism in the marine actinomycete genus *Salinispora*. *Proc Natl Acad Sci U S A.* 111(12):1130-1139.