

Bioinformatics tools for the discovery of new lipopeptides with biocontrol applications

M. Pupin · A. Flissi · P. Jacques · Valérie Leclère

Accepted: 6 July 2018

© Koninklijke Nederlandse Planteziektenkundige Vereniging 2018

Abstract As conventional or chemical pesticides have negative impact on environment and health of both farmer and consumers, it becomes relevant to develop alternative solutions to limit their use. In this context, innovative strategies to accelerate the development of biocontrol agents are welcome. For a decade of years, it has been demonstrated that lipopeptides are very efficient weapons against fungi responsible for crop diseases. Lipopeptides are secondary metabolites, produced by many microorganisms including beneficial rhizobacteria. The lipopeptide biosynthetic pathways include nonribosomal peptide synthetases. These modular enzymatic complexes work as assembly lines to build the peptides step by step, leading to the production of original peptide compounds with specific features as the presence of non proteinogenic monomers and cyclic and branched structures. In this paper, Florine and Norine bioinformatics tools, especially dedicated to non-ribosomal synthetases and their products are presented. Their use is mainly focused on the discovery of lipopeptides produced by *Bacillus* or *Pseudomonas* because they seem to represent a versatile reservoir of active secondary metabolites with promising activities for applications in phytosanitary area.

Keywords NRPS · Lipopeptide · Biocontrol · Bioinformatics · Norine

Introduction

Biocontrol (a contraction of “biological control”) is an alternative integrated pest management aimed to plant protection. As conventional or chemical pesticides have negative impact on environment and health of both farmer and consumers, it becomes relevant to develop alternative solutions to limit their use. Biocontrol agents (BCAs) are classified into 4 groups including natural substances, pheromones, macro-organisms and micro-organisms. Fungi in the genus *Trichoderma* have been used in biocontrol for about a century (Harman 2006). Bacterial microbes have also been shown to be well suited as BCAs. *Pseudomonas* and *Bacillus* species are the predominant plant growth promoting rhizobacteria (PGPR) known so far for these applications (Weller 2007; Radhakrishnan et al. 2017).

Fungal and bacterial biocontrol agents are especially efficient to control diseases caused by fungal and

M. Pupin · A. Flissi
Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRISTAL - Centre de Recherche en Informatique Signal et Automatique de Lille, F-59000 Lille, France

M. Pupin · A. Flissi
Inria-Lille Nord Europe, Bonsai Team, F-59655 Villeneuve d’Ascq Cedex, France

P. Jacques
TERRA Research Centre, Microbial Processes and Interactions (MiPI), Gembloux Agro-Bio Tech University of Liege, B-5030 Gembloux, Belgium

V. Leclère (✉)
Univ. Lille, INRA, ISA, Univ. Artois, Univ. Littoral Côte d’Opale, EA 7394-ICV- Institut Charles Viollette, F-59000 Lille, France
e-mail: Valerie.leclere@univ-lille1.fr

oomycete pathogens such as *Botrytis*, *Fusarium*, *Pythium*, *Phytophthora*, or *Rhizoctonia*, infecting crops of economic importance. They act as BCAs, through multiple direct and indirect mechanisms probably working in synergy to achieve disease control (Benítez et al. 2004; Harman 2006). Indeed, the final effect of BCAs may result from miscellaneous mechanisms including competition for nutrients and space, root colonization, modification of physico-chemical properties and microbiome of the rhizosphere, biofilm formation, and direct antibiosis related to the production of extracellular enzymes and secondary metabolites. Moreover, they can stimulate plant-defense mechanisms and increase plant growth. For *Bacillus* and *Pseudomonas* species, it has been demonstrated that the lipopeptides they produce play a major role in plant disease suppression (Ongena and Jacques 2008; Raaijmakers et al. 2006, 2010; Mnif and Ghribi 2015). As biocontrol exerted by fungal and bacterial living cells is sometimes unpredictable, the use of purified lipopeptides emerged, leading to the development of methods to screen for novel molecules. Bioassay-driven strategies can be used to detect antifungal activity (Mihalache et al. 2017), invasive behaviour (Béchet et al. 2013), or in vivo protective effect on plantlets (Leclère et al. 2005). These bio-assays are generally expensive and time consuming but the screening can be accelerated by PCR detection of lipopeptide synthetase genes (Tapi et al. 2010; Rokni-Zadeh et al. 2011). Another way leading to the identification of novel lipopeptides is high throughput bioinformatics screening of the numerous genome sequences available thanks to new sequencing techniques. In this review, the Florine (Caradec et al. 2014) and Norine (Flissi et al. 2016) bioinformatics tools aimed to speed up the in silico screening for novel lipopeptides are described. They may lead to the discovery of totally undescribed and untested lipopeptides or to the identification of novel variants belonging to known families, with differentiated activities usable in biocontrol strategies.

Lipopeptides as biocontrol agents

Microbial lipopeptides are secondary metabolites produced by a variety of fungi and bacteria, composed of a fatty acid hydrophobic tail linked to a peptide hydrophilic moiety. This amphiphilic structure explains their surface tension properties. The best biosurfactant known is surfactin secreted by *Bacillus subtilis* strains,

decreasing the surface tension of culture supernatant from 72 mN/m to 27 mN/m. Several hundreds of lipopeptides, grouped in about 40 families, are naturally produced by microorganisms. Applications in diverse area as health, agro-food or plant-protection depend on biological activities and properties related to structural features. Most of the active lipopeptides produced by plant beneficial bacteria are structurally related, cyclic or partially cyclic compounds referred as CLPs (cyclic lipopeptides) (Fig. 1). These compounds have versatile biological activities playing key roles in natural habitats. Thus, CLPs produced by *Pseudomonas* and *Bacillus* species exhibit growth inhibitory activity against a broad range of microorganisms including phytopathogenic fungi and oomycetes (Raaijmakers et al. 2006; Ongena and Jacques 2008; Raaijmakers et al. 2010). By changing the viscosity of the surface, CLPs are able to modify swarming and swimming behaviors (Leclère et al. 2006), increasing root colonization and biofilm formation (Raaijmakers et al. 2010; Chowdhury et al. 2015). More recently, some CLPs have been shown to trigger induced systemic resistance (ISR) and thus may contribute to biocontrol of diseases in crop (Ma et al. 2017).

Nonribosomal biosynthesis of lipopeptides

CLPs are biosynthesized by nonribosomal peptide synthetases (NRPSs). These enzymatic complexes are now considered as assembly lines, used as template and biosynthetic machinery. NRPS complexes may be constituted of one or more proteins, their genes being generally clustered on the genome, and for some of them constituting operons. Because NRPSs are modularly organized, they work step-by-step to assemble the building blocks (or monomers), leading to the construction of the peptides. Indeed, a module is defined as the part of the NRPS allowing the incorporation of one monomer (Fig. 2). For this purpose, modules are divided into enzymatic domains which are: i) adenylation domain (A) responsible for the selective recruitment of the monomers, ii) thiolation domain (T) covalently tethering the selected monomer and growing peptide chain on the NRPS, via a phosphopantetheynyl cofactor (the T-domain may also be annotated PCP for Peptidyl Carrier Protein); and iii) condensation domain (C) catalyzing the peptide bond formation. The global architecture of a module is thus (C-A-T). Ending the assembly line is the

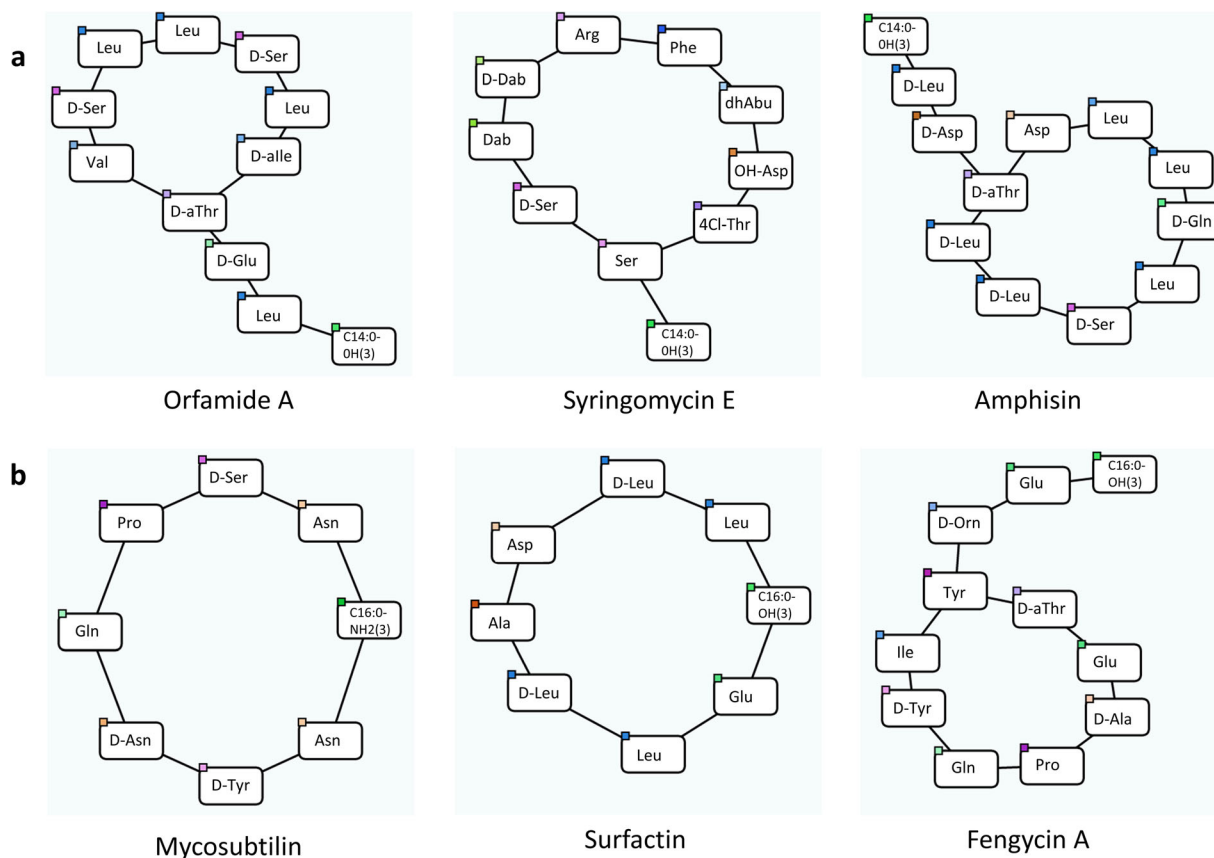


Fig. 1 Examples of CLPs produced by beneficial bacteria belonging to *Bacillus* and *Pseudomonas* genera. Monomeric representations are extracted from Norine platform. **a:** *Bacillus*, **b:** *Pseudomonas*

thioesterase domain (Te) releasing the neo-built peptide; for some of them the release is accompanied by cyclisation of the peptide (Martínez Núñez and Lopez 2016; Roongsawang et al. 2011; Strieker et al. 2010). CLP assembly lines of genera *Pseudomonas* and *Burkholderia* are frequently ended by a tandem of Te-domains (de Bruijn et al. 2007). Beside the core catalytic domains, secondary domains allowing the modification of the monomers during the synthesis can be present. The most frequently encountered are epimerization domains (E) modifying the L-form of the monomer into D-isomer, and to a lower extent methylation domains (M). While E-domains are located downstream T-domains, M-domains are present between A- and T-domains.

According to functional specificities, various subtypes of C-domains are described including $^L C_L$, $^D C_L$ (Rausch et al. 2007). With the exception of NRPS for lipopeptides belonging to the iturin family, the first domain of NRPS responsible for the synthesis of CLPs is a condensation domain called Cstarter, catalyzing the

condensation of the lipid tail on the first aminoacid of the peptide moiety (Rausch et al. 2007). First described in the arthrfactin synthetase (Balibar et al. 2005), dual C/E domains harboring both condensation and epimerization activities are usually present in all CLP NRPSs of *Pseudomonas* (Bloudoff and Schmeing 2017; Rausch et al. 2007; Ziemert et al. 2012). More recently, dual C/E domains were also identified within CLP NRPSs of soilborne rhizobacteria *Xanthomonas* (Royer et al. 2013) and *Burkholderia* (Esmaeel et al. 2016).

Bioinformatics screening for lipopeptides

Considering the specific features of both NRPS and their products, bioinformatics tools allowing the prediction of new non-ribosomal peptides from genomic data of microorganisms were developed. The Florine workflow proposes a sequential use of specific tools (Caradec et al. 2014), leading to the prediction of the

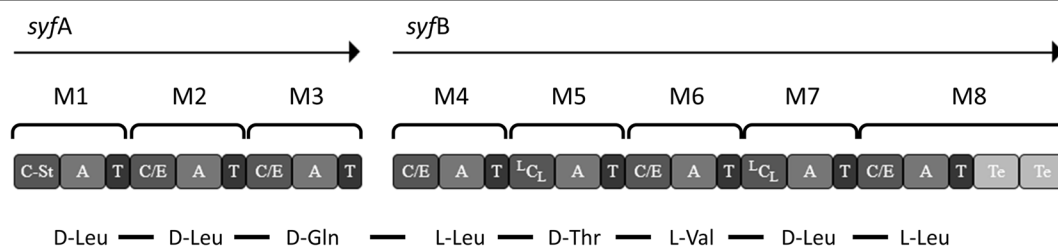


Fig. 2 Module and domain architecture of syringafactinA NRPS assembly line. Arrows represent the genes; M1 to M10 are the modules for incorporation of the 10 monomers; A: adenylation

domain; T: thiolation domain, C: condensation domains (with different sub-types C_{starter} , $^L C_L$, dualC/E); Te: thiolation domain; bottom: monomeric representation of syringafactinA

structure of non-ribosomal peptides potentially produced by a microorganism since the genome sequence is available (Fig. 3). In this workflow, the most popular and efficient tools are applied; each of them focusing on specific points to contribute to the global analysis and is allocated to the corresponding steps of the workflow. antiSMASH (Medema et al. 2011; Weber et al. 2015; Blin et al. 2013) is the most complete software to predict NRPS gene clusters and to analyze the modular and domain architecture of the enzymes. It is regularly updated making use of leading-edge tools. For example, SANDPUMA (Chevrette et al. 2017) published in 2017 has been integrated into antiSMASH4 (Blin et al. 2017). The first job of antiSMASH is the detection of gene clusters involved in nonribosomal biosynthesis. Thanks to specific domain signatures, the NRPS domains are

easily detected and the domain architecture of the proteins can be defined. Furthermore, in the last antiSMASH release (Blin et al. 2017) rules have been defined to detect clusters for various types of NRPs. It relies on comparative approaches to explore the annotations from public databases. However, no rule has been defined to automatically identify the product of a NRPS cluster as a lipopeptide. It is the reason why the design of a lipopeptide from the domain organization of the NRPS needs to be manually achieved. Attention has to be paid on both A- and C-domains. From A-domains, the most probably recruited monomer can be predicted by NRPSpredictor (Röttig et al. 2011). This prediction can be obtained directly from antiSMASH outputs which combine results from NRPSpredictor3, PrediCAT or SANDPUMA (Blin et al. 2017). Then, NaPDoS (Ziemert et al. 2012) that is also integrated in antiSMASH, facilitates C-domain sub-type determination. The identification of a C_{starter} in the first module of a synthetase leads to the prediction of a lipopeptide production. The D-isomery of the monomers can be predicted by the identification of $^D C_L$ domain following E-domain or identification of dual C/E-domain, especially for *Pseudomonas*, *Xanthomonas* or *Burkholderia* species. Finally, the assignation of a NRPS to a CLP synthetase can be strengthened by the identification of a tandem of Te-domains ending the assembly lines.

The novelty of a predicted CLP should be evaluated by structure comparison with all NRPs stored in Norine database (Caboche et al. 2008; Flissi et al. 2016). This step could output several cases as: i) the CLP exists, maybe produced by another microorganism; ii) the CLP is a new variant belonging to a known family; and iii) the CLP harbors a structure that is not described in Norine.

A direct link to query the Norine database with the predicted peptide structure is provided for this purpose on antiSMASH result page but antiSMASH results never consider the isomery of the monomers. For this

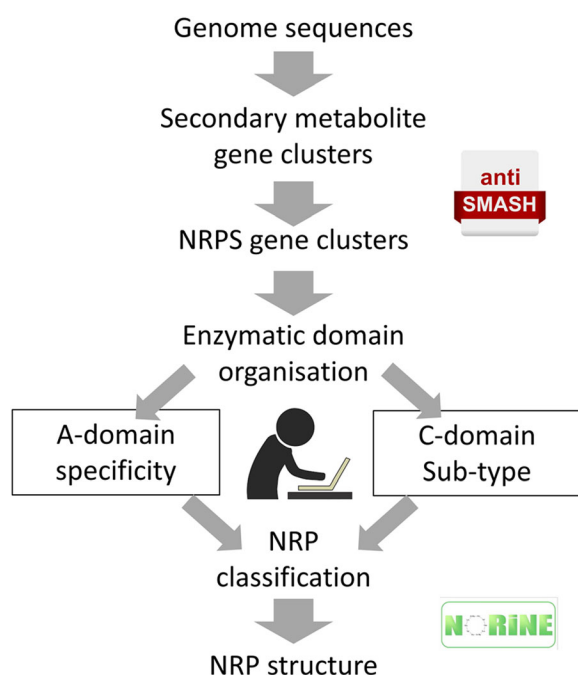


Fig. 3 Schematic representation of Florine workflow

purpose, the final step of the workflow performed using the Norine platform is essential.

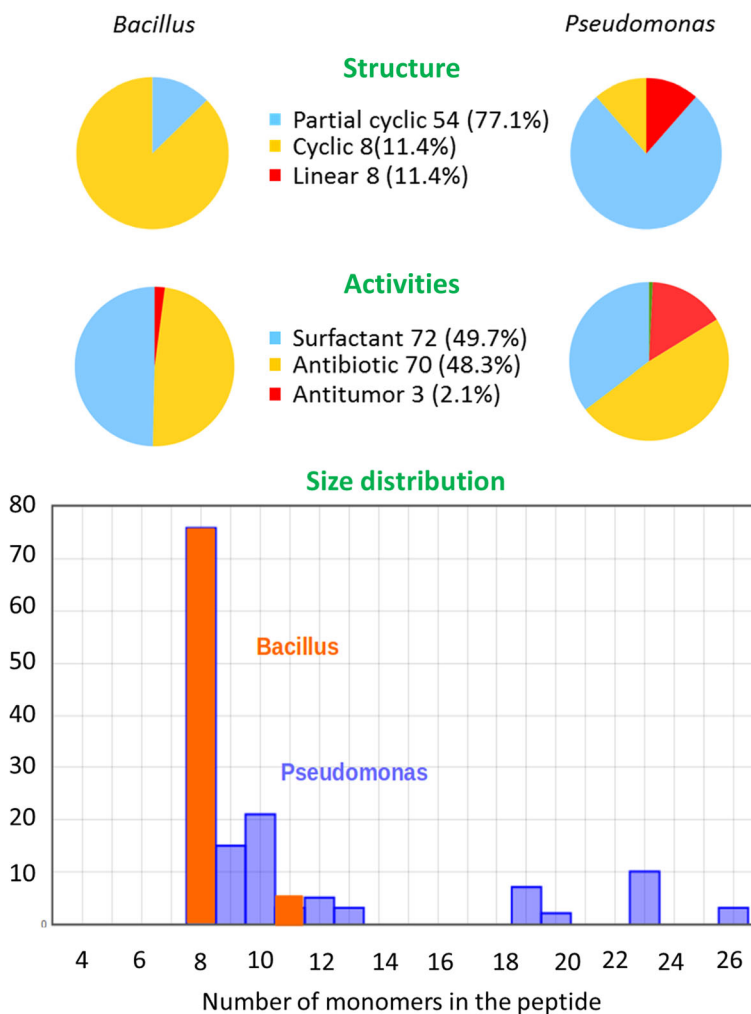
Norine platform dedicated to nonribosomal peptides

The Norine platform is composed of a database and different tools allowing biologists and biochemists to explore, visualize, query and analyze the NRPs. Since its creation a decade ago, Norine (Caboche et al. 2008) remains the unique most comprehensive knowledgebase dedicated to computational biology analysis of NRPs. The knowledgebase currently contains abundant and valuable annotations on about 1200 NRPs belonging to 217 families. These NRPs harbor original features mainly based on the wide diversity of monomers others than the 20 proteinogenic amino acids, and their complex structures.

Indeed, Norine counts about 530 monomers composing linear, cyclic, partial-cyclic and/or branched structures.

The Norine web application, available at <http://bioinfo.cristal.univ-lille.fr/norine> provides two kinds of interfaces to search for NRPs: annotation search and structure search. The first interface, named “annotation search”, allows querying the database by combining many criteria on NRPs such as the name, the category (e.g. glycopeptide, lipopeptide, etc.), the peptide molecular weight, the known activities (antibiotic, siderophore, surfactant, antiinflammatory...) or the producing organisms. The antibiotic activity includes antibacterial as well as antifungal because many compounds may have both. It is also possible to make a search using a title, journal or author name of the published references associated to the NRP. The main interest of the annotations search is to provide an overview on a set of NRPs that share

Fig. 4 Data downloaded from Norine for *Bacillus* and *Pseudomonas* NRP comparison



common properties, to explore and visualize them, analyze, refine results with filters and download data. Currently, Norine contains 1186 NRPs manually annotated. Among these NRPs, 299 are lipopeptides grouped in 36 families, such as orfamides and amphisines for *Pseudomonas* lipopeptides or surfactin and iturins for *Bacillus* lipopeptides. The NRP structure type is partial cyclic for half of them, then cyclic (37.6%) and linear (13.1%). Regarding the activities of these lipopeptides, it can be emphasized that antibiotic (223 peptides) and surfactant (143 peptides) activities are prevailing, some peptides displaying several activities. Finally, the size distribution shows that they contain between 4 and 26 monomers with a peak at 8 monomers. Using the annotations search interface of Norine the queries can be constructed in order to compare lipopeptides produced by *Bacillus* and *Pseudomonas* species (Fig. 4). For both genera, the structure type is mainly cyclic or partial cyclic. Indeed, more than 80% of them contains at least one cycle, the last monomer bridged with a monomer of the peptide chain or with the fatty acid chain for most of the *Bacillus* lipopeptides. Known activities of these lipopeptides are antibiotic (48.3% for *Bacillus*, 48.4% for *Pseudomonas*) and surfactant (respectively 49.7 and 34.4%), but also toxin (16.4% for *pseudomonas*) and antitumor. Concerning size distribution, i.e. the number of

monomers composing the lipopeptides, it may be of interest to point out that lipopeptides produced by *Bacillus* species are essentially composed of 8 monomers, whereas *Pseudomonas* species size are longer, between nine and up to 26 monomers (including the fatty acid moiety).

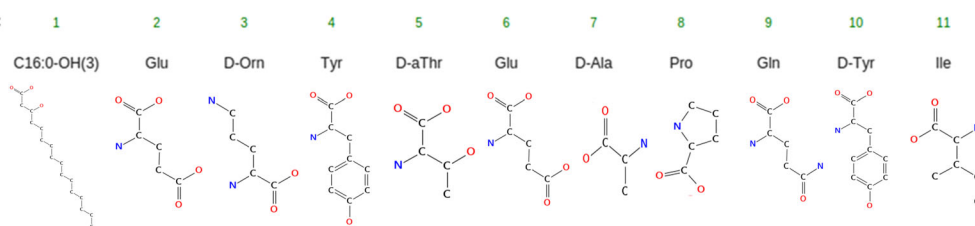
The second way to search for peptides is the “structure search” interface. This way is especially useful to assign novelty to NRPs, families or variants thanks to structure comparison. In Norine, each NRP has a 2D-monomeric representation of its structure, which is seen as a graph with nodes labeled by monomers. The complex 2D structures are represented using simple strings leading to the graph representation in Norine. As an example, the different representations are illustrated for the fengycin A lipopeptide (Fig. 5). The graph representation enables pattern research, i.e. search for sub-graph in a graph, with the intention of discovering potentially new NRPs. In addition, Norine provides a user-friendly NRP editor to draw parts of NRP and automatically transform it into graph representation for comparison (Fig. 6).

Finally, each NRP in Norine has a comprehensive result page gathering all the information about the peptide. This page is organized in six tabs: (i) general information on the peptide such as its name and category, formula or the known activities of the peptide; (ii) structure, including a visual 2D representation of the

▪ Smiles:

```
CCCCCCCCCCCCC(O)CC(=O)NC(CCC(=O)O)C(=O)NC(CCCN)C(=O)NC3Cc1ccc(cc1)OC(=O)C(NC(=O)C(Cc2ccc(O)cc2)NC(=O)C(CCC(N)=O)NC(=O)C4CCCN4(C(=O)C(C)NC(=O)C(CCC(=O)O)NC(=O)C(NC3(=O))C(C)O))C(C)CC
```

▪ Monomeric composition :



▪ Graph representation: C16:0-OH(3),Glu,D-Orn,Tyr,D-aThr,Glu,D-Ala,Pro,Gln,D-Tyr,Ile @1 @0,2 @1,3 @2,4,10 @3,5 @4,6 @5,7 @6,8 @7,9 @8,10 @3,9

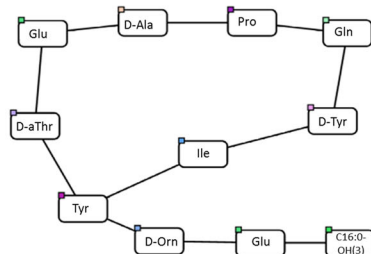


Fig. 5 Different representations of fengycinA lipopeptide as found in the Norine page

monomeric composition and peptide bounds (Fig. 5), and SMILES, the notation that represents the chemical structure of the peptide, if available; (iii) the producing organisms, (iv) a tree that represents the taxonomy of producing organisms; (v) up to two references describing the peptide structure, and its synthetase, when available; and (vi) finally, links to external databases such as PDB, PubChem, UniProt, StreptomeDB, ChEMBL or BIRD.

Thanks to “MyNorine” interface (Flissi et al. 2016) it is now easy for any scientist discovering a new NRP to enter it with annotations in the database and thus contribute to the enrichment of Norine. The increasing number of annotated NRPs is a key factor in the accuracy of structure comparison.

Conclusion

Identification of new NRPS, especially lipopeptides, is the more exciting and expected result of in silico screening for the discovery of novel biocontrol agents. Indeed, the bioinformatics workflow described here is useful to

detect cyclic lipopeptide synthesis gene clusters in a whole genome. However, as it remains only predictions, the next steps are experimental assays on the predicted compounds to confirm the structures and efficient production because many secondary metabolite genes may be cryptic. It is also necessary to verify if the potentially produced NRPs should actually be applicable as BCAs. Indeed, the Florine workflow has allowed the successful exploration of the genomes of 48 *Burkholderia* strains (Esmael et al. 2016) leading to the identification of a putative CLP named burkhomycin. The same CLP was independently described and named malleipeptin by Biggins et al. (2014), because the producing organism is *Burkholderia* belonging to the mallei group. Furthermore, malleipeptin was identified as a virulence factor for the producing bacterium, making its use as a BCA impossible. This example highlights that it is necessary to distinguish between the identification of a potentiality to produce secondary metabolites based on gene exploration in the whole genome and the direct use for diverse applications. Anyway, this screening remains interesting as it pointed out the presence of dual C/E domains in the

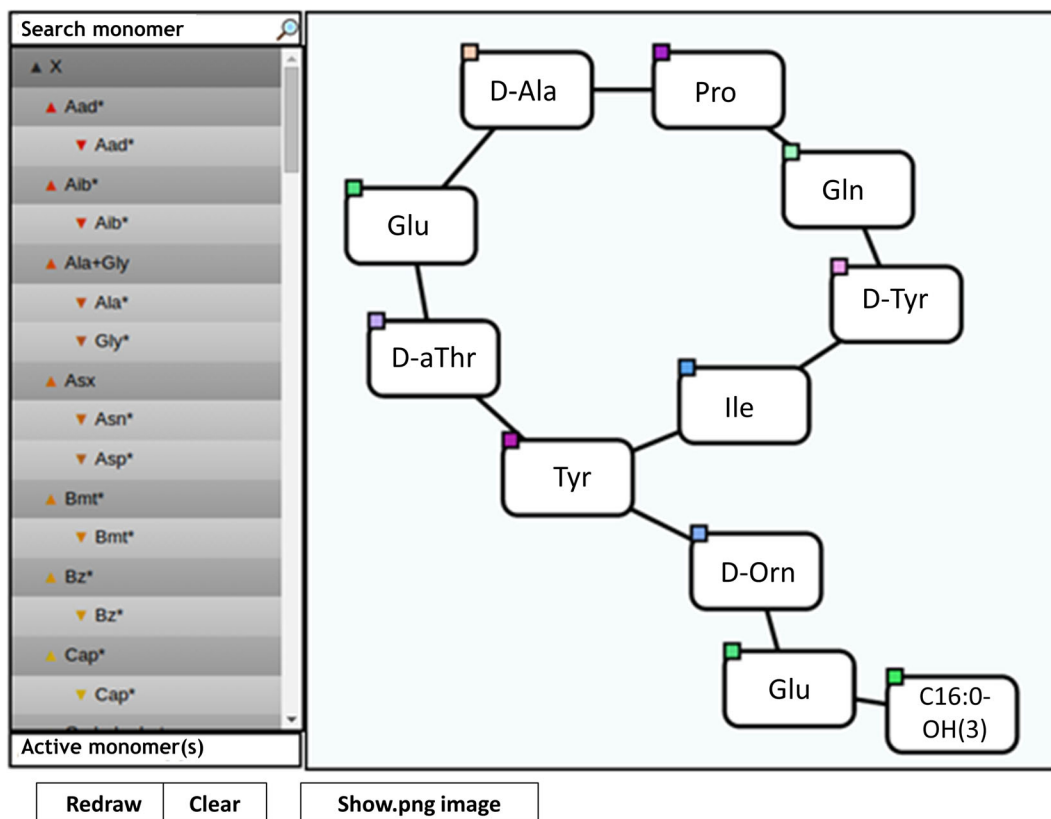


Fig. 6 Editor developed to transform automatically drawing into graph representation

CLP synthetase of *Burkholderia* strains. Elsewhere, the use of the dedicated bioinformatics tools, combined with biological assays led to successful identification of two CLPs produced by *Pseudomonas* sp. CMR12a (D'aes et al. 2014). The first CLP is orfamide B and the second one is a new member of the tolaasin family, named sessilin. Because of their biological activities in biofilm formation and because they are produced by the biocontrol strain CMR12a, both CLPs can be considered as good candidates for plant protection applications.

Acknowledgements This work has been carried out in the framework of Alibiotech project which is financed by European Union, French State and the French Region of Hauts-de-France. Authors would like to thank European Union funding through the INTERREG V France-Wallonie-Vlaanderen Project SmartBioControl/BioScreen and the bioinformatics platform bilille.

Compliance with ethical standards

Conflict of interest No Disclosure of potential conflicts of interest.

Human and animal rights No Research involving Human Participants and/or Animals.

Informed consent Yes

References

- Balibar, C. J., Vaillancourt, F. H., & Walsh, C. T. (2005). Generation of D amino acid residues in assembly of arthrofactin by dual condensation/epimerization domains. *Chemistry & Biology*, 12, 1189–1200. <https://doi.org/10.1016/j.chembiol.2005.08.010>.
- Béchet, M., Castéra-Guy, J., Guez, J.-S., et al. (2013). Production of a novel mixture of mycosubtilins by mutants of *Bacillus subtilis*. *Bioresource Technology*, 145, 264–270. <https://doi.org/10.1016/j.biortech.2013.03.123>.
- Benítez, T., Rincón, A. M., Limón, M. C., & Codón, A. C. (2004). Biocontrol mechanisms of *Trichoderma* strains. *International Microbiology*, 7, 249–260.
- Biggins, J. B., Kang, H.-S., Ternei, M. A., et al. (2014). The chemical arsenal of *Burkholderia pseudomallei* is essential for pathogenicity. *Journal of American Chemical Societies*, 136, 9484–9490. <https://doi.org/10.1021/ja504617n>.
- Blin, K., Medema, M. H., Kazempour, D., et al. (2013). antiSMASH 2.0—A versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Research*, 41, W204–W212. <https://doi.org/10.1093/nar/gkt449>.
- Blin, K., Kim, H. U., Medema, M. H., & Weber, T. (2017). Recent development of antiSMASH and other computational approaches to mine secondary metabolite biosynthetic gene clusters. *Brief Bioinformatics*. <https://doi.org/10.1093/bib/bbx146>.
- Bloudoff, K., & Schmeing, T. M. (2017). Structural and functional aspects of the nonribosomal peptide synthetase condensation domain superfamily: Discovery, dissection and diversity. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1865, 1587–1604. <https://doi.org/10.1016/j.bbapap.2017.05.010>.
- Caboche, S., Pupin, M., Leclère, V., et al. (2008). NORINE: A database of nonribosomal peptides. *Nucleic Acids Research*, 36, D326–D331.
- Caradec, T., Pupin, M., Vanvlassenbroeck, A., et al. (2014). Prediction of monomer isomery in Florine: A workflow dedicated to nonribosomal peptide discovery. *PLoS One*, 9, e85667. <https://doi.org/10.1371/journal.pone.0085667>.
- Chevrette, M. G., Aicheler, F., Kohlbacher, O., et al. (2017). SANDPUMA: Ensemble predictions of nonribosomal peptide chemistry reveal biosynthetic diversity across Actinobacteria. *Bioinformatics*, 33, 3202–3210. <https://doi.org/10.1093/bioinformatics/btx400>.
- Chowdhury, S. P., Hartmann, A., Gao, X., & Borriess, R. (2015). Biocontrol mechanism by root-associated *Bacillus amyloliquefaciens* FZB42 – A review. *Frontiers in Microbiology*, 6. <https://doi.org/10.3389/fmicb.2015.00780>.
- D'aes, J., Kieu, N. P., Leclère, V., et al. (2014). To settle or to move? The interplay between two classes of cyclic lipopeptides in the biocontrol strain *Pseudomonas* CMR12a. *Environmental Microbiology*, 16, 2282–2300. <https://doi.org/10.1111/1462-2920.12462>.
- de Bruijn, I., de Kock, M. J. D., Yang, M., et al. (2007). Genome-based discovery, structure prediction and functional analysis of cyclic lipopeptide antibiotics in *Pseudomonas* species. *Molecular Microbiology*, 63, 417–428. <https://doi.org/10.1111/j.1365-2958.2006.05525.x>.
- Esmael, Q., Pupin, M., Kieu, N. P., et al. (2016). *Burkholderia* genome mining for nonribosomal peptide synthetases reveals a great potential for novel siderophores and lipopeptides synthesis. *MicrobiologyOpen*, 5, 512–526. <https://doi.org/10.1002/mbo3.347>.
- Flissi, A., Dufresne, Y., Michalik, J., et al. (2016). Norine, the knowledgebase dedicated to non-ribosomal peptides, is now open to crowdsourcing. *Nucleic Acids Research*, 44, D1113–D1118. <https://doi.org/10.1093/nar/gkv1143>.
- Harman, G. E. (2006). Overview of mechanisms and uses of *Trichoderma* spp. *Phytopathology*, 96, 190–194. <https://doi.org/10.1094/PHYTO-96-0190>.
- Leclère, V., Béchet, M., Adam, A., et al. (2005). Mycosubtilin overproduction by *Bacillus subtilis* BBG100 enhances the organism's antagonistic and biocontrol activities. *Applied and Environmental Microbiology*, 71, 4577–4584. <https://doi.org/10.1128/AEM.71.8.4577-4584.2005>.
- Leclère, V., Marti, R., Béchet, M., et al. (2006). The lipopeptides mycosubtilin and surfactin enhance spreading of *Bacillus subtilis* strains by their surface-active properties. *Archives in Microbiology*, 186, 475–483. <https://doi.org/10.1007/s00203-006-0163-z>.
- Ma, Z., Ongena, M., & Höfte, M. (2017). The cyclic lipopeptide orfamide induces systemic resistance in rice to *Cochliobolus miyabeanus* but not to *Magnaporthe oryzae*. *Plant Cell*

- Reports*, 36, 1731–1746. <https://doi.org/10.1007/s00299-017-2187-z>.
- Martínez Núñez, M. A., & López y Lopez, V. E. L. (2016). Nonribosomal peptides synthetases and their applications in industry. *Sustain Chemistry Process*, 4, 13. <https://doi.org/10.1186/s40508-016-0057-6>.
- Medema, M. H., Blin, K., Cimermancic, P., et al. (2011). antiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Research*, 39, W339–W346. <https://doi.org/10.1093/nar/gkr466>.
- Mihalache, G., Balaes, T., Gostin, I., et al. (2017). Lipopeptides produced by *Bacillus subtilis* as new biocontrol products against fusariosis in ornamental plants. *Environmental Sciences and Pollution Research*, 1–10. <https://doi.org/10.1007/s11356-017-9162-7>.
- Mnif, I., & Ghribi, D. (2015). Microbial derived surface active compounds: Properties and screening concept. *World Journal of Microbiology Biotechnology*, 31, 1001–1020. <https://doi.org/10.1007/s11274-015-1866-6>.
- Ongena, M., & Jacques, P. (2008). *Bacillus* lipopeptides: Versatile weapons for plant disease biocontrol. *Trends in Microbiology*, 16, 115–125. <https://doi.org/10.1016/j.tim.2007.12.009>.
- Raaijmakers, J. M., de Bruijn, I., & de Kock, M. J. D. (2006). Cyclic lipopeptide production by plant-associated *Pseudomonas* spp.: Diversity, activity, biosynthesis, and regulation. *Molecular Plant-Microbe Interactions*, 19, 699–710. <https://doi.org/10.1094/MPMI-19-0699>.
- Raaijmakers, J. M., de Bruijn, I., Nybroe, O., & Ongena, M. (2010). Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: More than surfactants and antibiotics. *FEMS Microbiological Reviews*, 34, 1037–1062. <https://doi.org/10.1111/j.1574-6976.2010.00221.x>.
- Radhakrishnan, R., Hashem, A., & Abd Allah, E. F. (2017). *Bacillus*: A biological tool for crop improvement through bio-molecular changes in adverse environments. *Frontiers in Physiology*, 8, 667. <https://doi.org/10.3389/fphys.2017.00667>.
- Rausch, C., Hoof, I., Weber, T., et al. (2007). Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. *BMC Evolutionary Biology*, 7, 78. <https://doi.org/10.1186/1471-2148-7-78>.
- Rokni-Zadeh, H., Mangas-Losada, A., & Mot, R. D. (2011). PCR detection of novel non-ribosomal peptide synthetase genes in lipopeptide-producing *Pseudomonas*. *Microbial Ecology*, 62, 941. <https://doi.org/10.1007/s00248-011-9885-9>.
- Roongsawang, N., Washio, K., & Morikawa, M. (2011). Diversity of nonribosomal peptide synthetases involved in the biosynthesis of lipopeptide biosurfactants. *International Journal of Molecular Science*, 12, 141–172. <https://doi.org/10.3390/ijms12010141>.
- Röttig, M., Medema, M. H., Blin, K., et al. (2011). NRPSpredictor2—A web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Research*, 39, W362–W367. <https://doi.org/10.1093/nar/gkr323>.
- Royer, M., Koebnik, R., Marguerettaz, M., et al. (2013). Genome mining reveals the genus *Xanthomonas* to be a promising reservoir for new bioactive non-ribosomally synthesized peptides. *BMC Genomics*, 14, 658. <https://doi.org/10.1186/1471-2164-14-658>.
- Strieker, M., Tanović, A., & Marahiel, M. A. (2010). Nonribosomal peptide synthetases: Structures and dynamics. *Current Opinion in Structural Biology*, 20, 234–240. <https://doi.org/10.1016/j.sbi.2010.01.009>.
- Tapi, A., Chollet-Imbert, M., Scherens, B., & Jacques, P. (2010). New approach for the detection of non-ribosomal peptide synthetase genes in *Bacillus* strains by polymerase chain reaction. *Applied Microbiology Biotechnology*, 85, 1521–1531. <https://doi.org/10.1007/s00253-009-2176-4>.
- Weber, T., Blin, K., Duddela, S., et al. (2015). antiSMASH 3.0—A comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Research*, 43, W237–W243. <https://doi.org/10.1093/nar/gkv437>.
- Weller, D. M. (2007). *Pseudomonas* biocontrol agents of soilborne pathogens: Looking back over 30 years. *Phytopathology*, 97, 250–256. <https://doi.org/10.1094/PHYTO-97-2-0250>.
- Ziemert, N., Podell, S., Penn, K., et al. (2012). The natural product domain seeker NaPDoS: A phylogeny based bioinformatic tool to classify secondary metabolite gene diversity. *PLoS One*, 7, e34064. <https://doi.org/10.1371/journal.pone.0034064>.