



**HAL**  
open science

# Caractérisation chimique des métabolomes secondaires de *Penicillium* et *Fusarium* par marquage isotopique couplé à la spectrométrie de masse haute résolution

Thaïs Hautbergue

► **To cite this version:**

Thaïs Hautbergue. Caractérisation chimique des métabolomes secondaires de *Penicillium* et *Fusarium* par marquage isotopique couplé à la spectrométrie de masse haute résolution. Sciences agricoles. Institut National Polytechnique de Toulouse - INPT, 2017. Français. NNT: 2017INPT0110 . tel-04228503

**HAL Id: tel-04228503**

**<https://theses.hal.science/tel-04228503>**

Submitted on 4 Oct 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université  
de Toulouse

# THÈSE

En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

**Délivré par :**

Institut National Polytechnique de Toulouse (INP Toulouse)

**Discipline ou spécialité :**

Pathologie, Toxicologie, Génétique et Nutrition

---

**Présentée et soutenue par :**

Mme THAIS HAUTBERGUE

le mardi 14 novembre 2017

**Titre :**

Caractérisation chimique des métabolomes secondaires de *Penicillium* et *Fusarium* par marquage isotopique couplé à la spectrométrie de masse haute résolution

---

**Ecole doctorale :**

Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries (SEVAB)

**Unité de recherche :**

Toxicologie Alimentaire (ToxAlim)

**Directeur(s) de Thèse :**

MME ISABELLE OSWALD

M. EMILIE JAMIN

**Rapporteurs :**

M. DAVID TOUBOUL, CNRS GIF SUR YVETTE

M. JOSE DIANA DI MAVUNGU, UNIVERSITE DE GENT

**Membre(s) du jury :**

Mme FLORENCE MATHIEU, INP TOULOUSE, Président

M. CEDRIC BERTRAND, UNIVERSITE DE PERPIGNAN, Membre

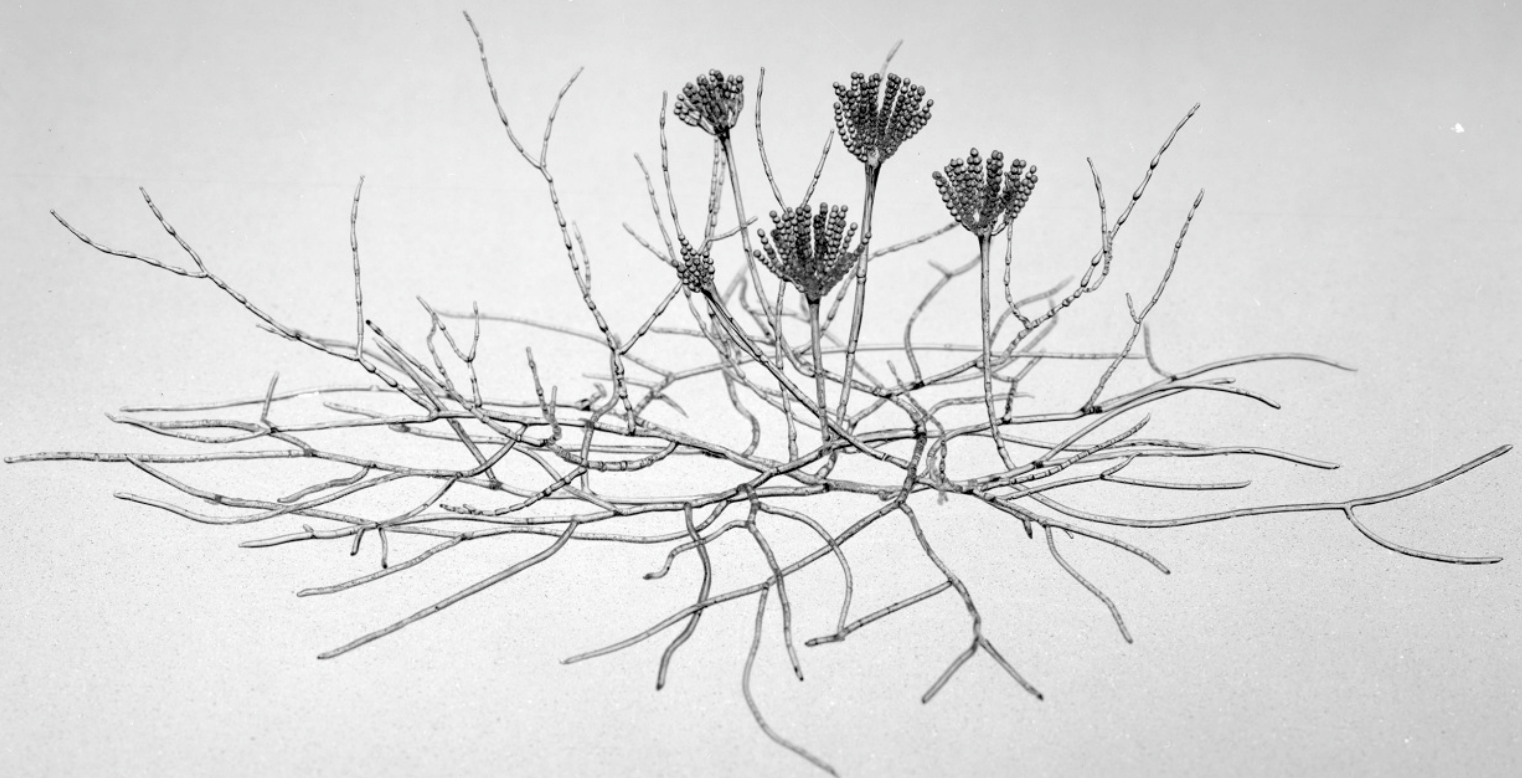
M. FABIEN JOURDAN, INRA TOULOUSE, Membre

M. OLIVIER PUEL, INRA TOULOUSE, Membre



« Le rôle des infiniment petits dans la nature  
est infiniment grand »

Louis Pasteur



À Élise



## REMERCIEMENTS

Je remercie le Ministère de l'Éducation nationale, de l'Enseignement Supérieur et de la Recherche, les porteurs du projet MetaboHUB-ANR-11-INBS-0010, l'ensemble des représentants de l'INRA et en particulier de l'unité Toxalim pour leur soutien matériel et financier sans lesquels je n'aurais pas pu mener cette thèse à son terme. Je remercie également les représentants de la plateforme MetaboHub-MetaToul de m'avoir intégrée dans leur structure ainsi que l'INP et l'école doctorale SEVAB. Parmi eux, je remercie tout particulièrement Claude Maranges pour l'attention et la bienveillance dont il fait preuve envers les doctorants de l'école doctorale SEVAB.

Je tiens à remercier Isabelle Oswald et Emilien Jamin qui ont accepté de diriger ma thèse, et bien sûr Olivier Puel pour avoir été mon troisième encadrant majeur. Merci Isabelle pour m'avoir accueillie au sein de votre équipe et pour m'avoir fait confiance au cours de cette thèse. Je vous remercie d'avoir été le chef d'orchestre nécessaire au bon fonctionnement de cette thèse ainsi que pour votre importante réactivité malgré les nombreuses responsabilités qui vous incombent. Emilien, merci également de m'avoir formée sur les techniques analytiques, pour ta réactivité et pour ton implication particulière dans ce projet. Merci de m'avoir fait confiance, de m'avoir accordé une grande autonomie sur ce projet, tout en étant bien sûr présent à chacune des étapes. De nombreux paramètres (aspect spectrométrie de masse de ma thèse, le fait que tu sois mon directeur de thèse, que je sois ta première thésarde, cohabitation au sein du bâtiment A) ont fait que tu as été l'encadrant avec lequel j'ai le plus échangé au cours de ces trois années. Tu as notamment veillé à ce que j'apprenne à canaliser mon énergie et à ce que je découvre tous les aspects de la vie en entreprise. Olivier, merci pour ton expertise, mais aussi pour ton optimisme qui m'a plus d'une fois réconfortée dans mes moments de doute. Merci beaucoup pour ton implication, pour avoir supervisé l'aspect fongique de ma thèse et pour m'avoir accordé le temps dont j'avais besoin malgré tes nombreux projets.

Je voudrais également remercier les membres de mon comité de thèse. Laurent Debrauwer, la force tranquille qui veille de loin sur chacun des membres de son équipe. Merci pour votre humanité et votre gentillesse qui m'ont permis de garder la tête hors de l'eau dans les moments difficiles. Cette énergie pour continuer malgré les difficultés, je l'ai également puisée chez Jean-Claude Tabet. Les semaines intensives d'interprétation de nos spectres et de réflexion scientifique m'ont permis de vivre pleinement ma passion pour ces casse-têtes et de m'épanouir professionnellement. J'ai rencontré au début de ma thèse un véritable mentor, un exemple autant sur le plan technique qu'humain. Merci pour tout Jean-Claude, de la part de tous tes étudiants, sincèrement. Enfin, merci Yann Guitton pour ton expertise en bioinformatique. Tu as cru dans ce projet et dans ses perspectives. Ton regard extérieur toujours très positif n'a fait que renforcer ma motivation.

Je souhaite remercier les rapporteurs de ma thèse, David Touboul et José Diana di Mavungu d'avoir accepté d'accorder de leur précieux temps à la lecture de mon manuscrit, ainsi que les membres du jury pour avoir accepté d'être présents le jour de ma soutenance. Merci à vous tous de le faire part de votre expertise.

Je tiens à remercier également tous les membres d'AXIOM pour leur aide tout au long de ce projet. Merci à l'équipe 5 de m'avoir accueillie chaleureusement. Merci Soraya pour tes précieux conseils, merci Sylviane et Jean-Denis pour votre bienveillance et votre aide. Je dédie une pensée affective pour tous les « jeunes » du labo grâce à qui la thèse représente également de super souvenirs amicaux. Ceux du bureau, tout d'abord, avec qui j'ai partagé notre « placard ». Aurélien (désolée encore de t'avoir rendu la vie aussi difficile !), Christelle, Djawed, Adéline, Salim, Robin et tout particulièrement Claire notre petite sœur si présente. Alyssa, on est arrivée en même temps mais toi, tu as pris perpet' comme on dit. Selma, Maria, Clément, Cyndel, Jade, Elodie, Morgane, Laura, Perrine, mais aussi tous les geeks, les jeunes de MeX et les filles de l'équipe 5, merci. Je reste fière d'avoir supervisé la première et sûrement dernière maison de campagne construite à l'intérieur des bureaux de l'INRA. La glycine n'a jamais aussi bien poussé que dans notre cagibi. Enfin, merci à tous ceux qui m'ont entourée de près ou de loin dans les moments difficiles et qui ont stimulé mes désirs de réussite.

Obrigada à minha revisora favorito. Saúdo sua coragem nesta tarefa ingrata.



# TABLE DES MATIÈRES

<b>COMMUNICATIONS SCIENTIFIQUES</b> .....	i
<b>TABLE DES ILLUSTRATIONS</b> .....	iii
<b>LISTE DES ABBRÉVIATIONS</b> .....	v
<b>INTRODUCTION</b> .....	1
1. CONTEXTE DE L'ÉTUDE.....	3
2. OBJECTIFS ET DÉROULEMENT DE LA THÈSE.....	5
3. ARTICLE 1 : CARACTÉRISATION DES MÉTABOLITES SECONDAIRES FONGIQUES.....	7
4. DES ESPÈCES FONGIQUES TOXINOGENES.....	33
4.1. <i>P. verrucosum</i> et <i>P. nordicum</i> .....	33
4.1.1. Le genre <i>Penicillium</i> .....	33
4.1.2. Propriétés de <i>P. verrucosum</i> et <i>P. nordicum</i> .....	34
4.1.3. Critères de classification de <i>P. verrucosum</i> et <i>P. nordicum</i> .....	35
4.1.4. Métabolomes secondaires de <i>P. verrucosum</i> et <i>P. nordicum</i> .....	37
4.2. <i>F. graminearum</i> , un champignon phytopathogène prédominant en Europe.....	39
4.2.1. Présentation de <i>F. graminearum</i> .....	39
4.2.2. Le métabolome secondaire de <i>F. graminearum</i> .....	41
5. LES FUNGISPORINES.....	42
5.1. Une découverte ancienne, une caractérisation récente.....	42
5.2. Origines génétiques des fungisporines.....	44
5.3. Des métabolites d'intérêt.....	45
6. RÉFÉRENCES DE L'INTRODUCTION.....	47
<b>MATÉRIELS &amp; MÉTHODES</b> .....	49
1. LA GÉNÉRATION DES DONNÉES.....	51
1.1. La méthode analytique.....	51
1.1.1. Le principe.....	51
1.1.2. La métabolomique non ciblée.....	52
1.1.3. Marquage isotopique du métabolome secondaire fongique.....	52
1.1.4. Analyse des métabolomes par HPLC-HRMS.....	53
1.2. Analyse des données.....	54
1.2.1. Extraction des pics.....	54
1.2.2. Suppression de la redondance d'information.....	57
1.2.3. Détermination des formules chimiques.....	58
1.3. Analyse des métabolites secondaires par MS/MS.....	61
1.3.1. Génération des spectres MS/MS par acquisition data dépendante.....	61
1.3.2. Analyse ciblée des métabolites secondaires fongiques en mode CID.....	61
1.3.3. Analyse ciblée des métabolites secondaires fongiques en mode HCD.....	62
2. INTERPRÉTATION DES RÉSULTATS.....	62
2.1. Les bases de données.....	62
2.2. Validation des annotations : identification.....	62
2.3. Les réseaux moléculaires : Le "GNPS Molecular Networking System".....	63
3. LE SÉQUENÇAGE PEPTIDIQUE <i>DE NOVO</i> .....	64
3.1. Les outils de séquençage peptidique automatique.....	64
3.2. Le séquençage peptidique <i>de novo</i> .....	65
3.2.1. Le modèle du proton mobile.....	65
3.2.2. Étude théorique des mécanismes de fragmentation des peptides linéaires.....	65
3.2.3. Les mécanismes de fragmentation des adduits sodium.....	70
3.2.4. Les mécanismes de fragmentation des peptides cycliques.....	72
4. RÉFÉRENCES DU CHAPITRE MATÉRIELS & MÉTHODES.....	75



<b>TRAVAIL EXPÉRIMENTAL</b> .....	77
<b>PARTIE 1. CARACTÉRISATION DU MÉTABOLOME SECONDAIRE DE <i>P. verrucosum</i></b> .....	79
1. ARTICLE 2. CARACTÉRISATION DU MÉTABOLOME SECONDAIRE DE <i>P. verrucosum</i> .....	81
2. OPTIMISATION DES RÉSEAUX MOLÉCULAIRES .....	102
2.1. Optimisation des données.....	102
2.2. Optimisation des paramètres du réseau moléculaire .....	102
2.3. Visualisation du réseau moléculaire .....	105
3. INTERPRÉTATION DES RÉSEAUX MOLÉCULAIRES .....	105
4. RÉFÉRENCES QUANT À L'ÉTUDE DE <i>P. verrucosum</i> .....	107
<b>PARTIE 2. LA PRODUCTION DE FUNGISPORINES PAR <i>P. nordicum</i></b> .....	109
1. ARTICLE 3. RECHERCHE DE NOUVEAUX PEPTIDES CHEZ <i>P. nordicum</i> .....	111
2. ARTICLE 4. UN MÉCANISME DE FRAGMENTATION PEPTIDIQUE NON RÉFÉRENCÉ .....	137
<b>PARTIE 3. ÉTUDE DU MÉTABOLOME SECONDAIRE DE <i>F. graminearum</i></b> .....	147
1. ARTICLE 5. LE MÉTABOLOME D'UN PUISSANT PHYTOPATHOGENÈ.....	149
<b>DISCUSSION</b> .....	173
1. ANALYSE CRITIQUE DE LA MÉTHODE .....	177
1.1. Marquage isotopique du substrat fongique.....	177
1.1.1. Une méthode adaptable à de multiples sujets d'étude .....	177
1.1.2. Limites de la méthode .....	177
1.2. Les outils bioinformatiques .....	178
1.2.1. CFM-ID .....	178
1.2.2. Les réseaux moléculaires .....	179
1.2.3. Développement d'un nouvel outil bioinformatique.....	180
1.3. Analyse critique de l'étude métabolomique des champignons filamenteux.....	181
1.4. Discussion autour de la fragmentation peptidique.....	183
2. ANALYSE DES RÉSULTATS DE L'ÉTUDE .....	184
2.1. Environ 80% de métabolites secondaires inconnus .....	184
2.2. Le métabolome secondaire de <i>F. graminearum</i> .....	185
2.2.1. Des métabolomes secondaires différents selon la souche et l'environnement.....	185
2.2.2. Découvertes de nouvelles fusaristatines chez <i>F. graminearum</i> .....	186
2.3. Les métabolomes secondaires des espèces de <i>Penicillium</i> productrices d'ochratoxine A .....	189
2.3.1. Caractérisation de métabolites secondaires connus.....	189
2.3.2. Découverte d'un nouveau peptide cyclique.....	190
2.3.3. Caractérisation de nouveaux térapeptides linéaires.....	192
2.3.4. Production de nouveaux pentapeptides par <i>P. nordicum</i> .....	194
2.3.5. Hypothèses quant aux propriétés des térapeptides produits les <i>Penicillium</i> .....	194
2.3.6. Perspectives pour l'analyse de ces composés .....	195
3. CLASSIFICATION MÉTABOLIQUE ET NOTION D'ESPÈCE FONGIQUE .....	197
3.1. Proximité phylogénique et similarité métabolique.....	197
3.2. Une souche de <i>P. verrucosum</i> chémotype anacine .....	197
3.3. Réflexion autour de la notion de caractérisation métabolique .....	198
3.3.1. Mise au point du système.....	198
3.3.2. Identification d'une souche grâce à ce système.....	200
3.3.3. Analyse critique d'un tel système de classification métabolique.....	200
4. RÉFÉRENCES DE LA DISCUSSION .....	202
<b>CONCLUSION</b> .....	203

## COMMUNICATIONS SCIENTIFIQUES

### ARTICLES

**T. Hautbergue**, O. Puel, S. Tadrict, L. Meneghetti, M. Péan, M. Delaforge, L. Debrauwer, I. P. Oswald and E. L. Jamin « *Evidencing 98 secondary metabolites of *Penicillium verrucosum* using substrate isotopic labeling and high resolution mass spectrometry* », *Journal of Chromatography B*. Sous presse. doi: 10.1016/j.jchromb.2017.03.011

**T. Hautbergue**, O. Puel, L. Debrauwer, E. L. Jamin and I. Oswald « *From genomic to metabolomic, moving toward an integrated strategy for the discovery of fungal secondary metabolites* », *soumis à Natural Product Reports*

**T. Hautbergue et al.** « *Ochratoxin A producing *Penicillium* species: new tetrapeptides on *Penicillium nordicum*'s secondary metabolome* », en préparation

**T. Hautbergue et al.** « *Evidencing a new peptidic fragmentation mechanism* », en préparation

P. Cano, E. Jamin, S. Tadrict, **T. Hautbergue**, P. Bourdaudhui, M. Péan, L. Debrauwer, I. P. Oswald, M. Delaforge, O. Puel « *Identification of new secondary metabolites produced in vitro and in planta by *F. graminearum* on wheat using isotopic labeling and HPLC-HRMS* », en préparation

### CONFÉRENCES

**T. Hautbergue**, O. Puel, S. Tadrict, M. Péan, M. Delaforge, L. Debrauwer, I. P. Oswald, E. L. Jamin « *Identification of six new fungisporins in *Penicillium nordicum* by stable isotope labeling and HPLC-HRMS* » 33<sup>e</sup> Journées de Spectrométrie de Masse (JFSM), 2016, Bordeaux, France.

**T. Hautbergue**, O. Puel, S. Tadrict, M. Péan, M. Delaforge, L. Debrauwer, I. P. Oswald, E. L. Jamin « *Molecular networks: optimization of parameters to study the secondary metabolome of *Penicillium verrucosum* and *Penicillium nordicum** » Rencontre du Club des Jeunes en Spectrométrie de Masse, 2016, Moustier-Sainte-Marie, France.

#### Prix de la meilleure présentation orale

**T. Hautbergue**, O. Puel, S. Tadrict, M. Péan, M. Delaforge, L. Debrauwer, I. P. Oswald, E. L. Jamin « *Evidencing 161 new secondary metabolites in *Penicillium* by stable isotope labeling and HPLC-HRMS* » Journée Nationale des Mycotoxines, 2016, Toulouse, France.

**T. Hautbergue**, O. Puel, S. Tadrict, M. Péan, M. Delaforge, L. Debrauwer, I. P. Oswald, E. L. Jamin « *Double stable isotope labeling and high-resolution mass spectrometry for the characterization of fungal secondary metabolomes* » Journée de Spectrométrie de Masse en Midi-Pyrénées (SMMP), 2015, Toulouse, France.

S. Chevolleau, I. Jouanin, **T. Hautbergue**, O. Martin, F. Guéraud, F. Pierre, L. Debrauwer « *Development of a LC-HRMS method for the non-targeted analysis of aldehydes: application for the detection of peroxidation products in sewage waters* » 10<sup>th</sup> Congress of the Francophone Association Francophone des Sciences Séparatives (AFSEP), 2013, Paris, France.

### POSTERS

**T. Hautbergue**, E. L. Jamin, O. Puel, L. Debrauwer and I. P. Oswald « *Characterization of the secondary metabolome of *Penicillium* by stable isotope labeling and HPLC HRMS* » 65<sup>e</sup> Congrès de l'American Society for Mass Spectrometry (ASMS), 2016, Indianapolis, USA.

**T. Hautbergue**, E. L. Jamin, C. Canlet, S. Tadrict, P. Bourdaudhui, O. Puel, L. Debrauwer and I. P. Oswald « *Characterization of new secondary metabolites of *Penicillium* using isotopic labeling and HPLC-HRMS* » 21<sup>e</sup> International Mass Spectrometry Conference (IMSC), 2016, Toronto, Canada.

**T. Hautbergue**, O. Puel, S. Tadrict, M. Péan, M. Delaforge, L. Debrauwer, I. P. Oswald, E. L. Jamin « *Characterization of the secondary metabolome of Penicillium by stable isotope labeling and HPLC HRMS* » Congrès de Spectrométrie de Masse et d'Analyse Protéomique (SMAP), 2015, Ajaccio, France.

**T. Hautbergue**, **E. L. Jamin**, C. Canlet, S. Tadrict, P. Bourdaudhui, O. Puel, L. Debrauwer and I. P. Oswald « *Double stable isotope labeling and high-resolution mass spectrometry for the characterization of fungal secondary metabolomes* » 11<sup>e</sup> Annual International Conference of the Metabolomics Society, 2015, San Francisco, USA.

## MÉDIATION SCIENTIFIQUE

- **Article Web pour l'Institut National de Recherche Agronomique (INRA)**

**Titre** : « Un champignon passé au crible des chercheurs : 86 nouvelles molécules chez *Penicillium verrucosum* »

**Lien** : <http://www.sa.inra.fr/Toutes-les-actualites/Un-champignon-passe-au-crible-des-chercheurs-86-nouvelles-molecules-chez-Penicillium-verrucosum>

- **Concours de médiation scientifique « Ma thèse en 180 secondes », France**

**Lien vers la vidéo de la prestation en finale nationale** : <https://www.youtube.com/watch?v=JxQY-0N1Sgk>

**Règles** : Les participants ont 180 secondes pour expliquer le sujet de leur thèse en français et en termes simples, à un auditoire profane et diversifié. Ils sont évalués par des représentants du monde socio-économique et du monde médiatique ainsi que des acteurs de la recherche sur le contenu et la clarté de leur présentation mais également sur leur charisme.

**Étapes de sélection** : Mars 2016 Premier prix du jury et premier prix du public en finale régionale  
Mai 2016 Première place en demie finale nationale

### Conférences invitées dans le cadre de MT180 :

Juin 2016 Présentation pour BioMedical Alliance et Bioméridies (Narbonne, France)

Janvier 2017 Présentation lors des vœux de la Ministre de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche Najat Vallaud-Belkacem (Musée de l'Homme, Paris, France)



## TABLE DES ILLUSTRATIONS

- Figure 1 Présentation des travaux expérimentaux réalisés aux cours de la thèse
- Figure 2 Classification des *Penicillium* selon le degré de ramification de leur mycélium.
- Figure 3 *P. verrucosum* : conidiophores et conidies.
- Figure 4 Classification phylogénique de *P. verrucosum* et *P. nordicum* au sein du genre *Penicillium*.
- Figure 5 Culture sur milieu YES agar de *P. verrucosum* et *P. nordicum* après 7 jours à 25°C.
- Figure 6 Métabolites secondaires de *P. verrucosum* et *P. nordicum* recensés dans la littérature.
- Figure 7 Observation macroscopique et microscopique de *F. graminearum*.
- Figure 8 Épi de blé infesté par *F. graminearum*.
- Figure 9 Principaux métabolites secondaires de *F. graminearum*.
- Figure 10 Lien hypothétique entre les tétrapeptides linéaires et les tétrapeptides fongiques.
- Figure 11 Schématisation des gènes *hcpA* d'*A. niger* et de *P. chrysogenum* responsables notamment de la synthèse de fungisporines chez ces champignons.
- Figure 12 Culture de *P. chrysogenum*. Souche sauvage et mutant  $\Delta hcpA$ .
- Figure 13 Stratégie analytique mise en œuvre pour l'analyse des métabolomes secondaires fongiques par double marquage isotopique et spectrométrie de masse haute résolution.
- Figure 14 Cibles d'un marquage isotopique selon son degré de spécificité du précurseur marqué.
- Figure 15 Schéma d'un spectromètre de masse LTQ-Orbitrap.
- Figure 16 Protocole d'extraction des pics.
- Figure 17 Extrait du spectre de masse issu de l'analyse du métabolome secondaire de *P. verrucosum* cultivé sur blé  $^{13}C^{15}N$ .
- Figure 18 Superposition des massifs isotopiques observés sur les spectres de masse d'un même métabolite marqué 53%  $^{13}C$  97%  $^{15}N$  (A), 97%  $^{13}C$  (B) et 99%  $^{12}C$  (C).
- Figure 19 Comparaison des spectres MS/MS obtenus par acquisition en mode data-dépendant des trois ions moléculaires détectés dans les spectres de masse  $^{12}C$ ,  $^{13}C$  et  $^{13}C^{15}N$ , d'un métabolite équivalent.
- Figure 20 Utilisation du logiciel MassToFormulaCompare.
- Figure 21 Lexique relatif aux réseaux.
- Figure 22 Nomenclature de la fragmentation peptidique en spectrométrie de masse.
- Figure 23 Fragmentation peptidique par formation d'une oxazolone et dissociations consécutives.
- Figure 24 Fragmentation peptidique générée par formation d'une dicétopipérazine.
- Figure 25 Formation des ions  $\mathbf{y}$  par rupture hétérolytique ou par formation d'une aziridinone.
- Figure 26 Perte d'eau à partir de la fonction acide carboxylique terminale.
- Figure 27 Formation d'ions immonium internes à partir d'ions  $\mathbf{y}$ .
- Figure 28 Hypothèses d'attaque nucléophile expliquant la perte de  $NH_3$  à partir de la fonction amine terminale d'un tryptophane.
- Figure 29 Mécanisme de fragmentation peptidique des adduits sodium d'après Renner *et al.*
- Figure 30 Interaction d'un atome de sodium avec les diverses fonctions d'un tétrapeptide.
- Figure 31 Mécanismes de fragmentation des peptides cycliques protonés en mode CID de basse énergie.
- Figure 32 Mécanismes de fragmentation des peptides cycliques adduits au sodium.
- Figure 33 Complexation du sodium sous forme de zwitterion et fragmentation des peptidiques cycliques adduits au sodium.
- Figure 34 Optimisation des paramètres de GNPS en fonction de la diversité des spectres MS/MS à analyser.
- Figure 35 Comparaison des spectres MS/MS de l'ochratoxine A, de l'ochratoxine B et du métabolite  $C_{17}H_{14}O_5$  détecté à 11,5 min.
- Figure 36 Protocole d'extraction et de traitement des données de l'outil en développement.
- Figure 37 Utilisation de l'échange hydrogène/deutérium pour la validation d'hypothèses structurales.
- Figure 38 Voies de biosynthèse de la fusaristatine A.

- Figure 39 Hypothèses quant à l'origine génétique de la synthèse de fusaristatine C.  
Figure 40 Schématisation d'un système de classification potentiel permettant la classification des espèces fongiques selon leur métabolome secondaire.

### LISTE DES TABLEAUX

- Tableau 1 Critères de classification de *P. verrucosum* et *P. nordicum*.  
Tableau 2 Tétrapeptides cycliques, tétrapeptides linéaires non modifiés et autres composés associés référencés dans la littérature.  
Tableau 3 Écarts de masse théoriques entre l'ion moléculaire ou le dimère et les adduits correspondants.  
Tableau 4 Exemple du calcul de la composition en carbone et en azote de l'ion détecté à  $m/z_{12C}=553.2814$  et détermination de sa formule brute.  
Tableau 5 Informations relatives aux 20 acides aminés protéinogènes standards.  
Tableau 6 Paramètres de génération de réseaux moléculaires par l'outil du GNPS.  
Tableau 7 Analyse des similarités entre les sommets annotés et les sommets adjacents inconnus du réseau moléculaire de *P. verrucosum*.  
Tableau 8 Peptides détectés chez *P. verrucosum* et *P. nordicum*.

## LISTE DES ABRÉVIATIONS

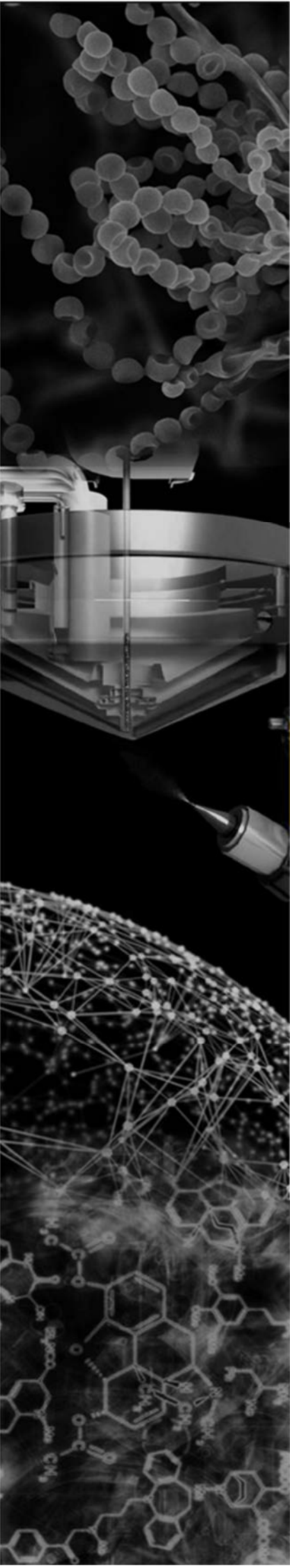
### *Abréviations en anglais*

CID	Collision-induced dissociation
DDA	Data dependent acquisition
ECD	Electron-capture dissociation
ESI	Electrospray ionization
ETD	Electron-transfert dissociation
HCD	Higher-energy collisional dissociation
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
IC50	Half-maximal inhibitory concentration
ITS	Internal transcribed spacer
LC/MS	Liquid chromatography coupled to mass spectrometry
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NRPS	Non-ribosomal peptide synthetase
OSMAC	One strain many compounds
PDA	Potato dextrose agar
PKS	Polyketides Synthase
TIC	Total ion current
USA	United states of america
YES	Yeast extract sucrose

### *Abréviations en français*

ADN	Acide désoxyribonucléique
CEA	Commissariat à l'énergie atomique et aux énergies alternatives
IR	infrarouge
$m/z$	Rapport masse sur charge
OTA	Ochratoxine A
ppm	Partie par million (e.g. mg/kg)
RMN	Résonance magnétique nucléaire
TR	Temps de rétention
UE/EU	Union européenne/European union
UV	Ultra-violet





# Introduction

« Which is better: to have fun with fungi or to have tomorrow's misdeeds out of yesterday's miscreeds? »

*Aldous Huxley*





### 1. CONTEXTE DE L'ÉTUDE

---

Les moisissures sont des microorganismes eucaryotes appartenant au règne fongique et dotés d'un développement filamenteux. Selon leur pouvoir pathogène et le milieu qu'ils colonisent, ces champignons microscopiques peuvent être impliqués dans des problématiques très diverses. Certaines espèces sont notamment responsables d'altérations directes de la santé humaine appelées mycoses. Alors que les dermatophytoses localisées sont pour la plupart bénignes, les mycoses systémiques affectant les individus immunodéprimés entraînent 1,6 à 2 millions de décès chaque année dans le monde<sup>1</sup>. Les genres *Candida* et *Aspergillus* comptent parmi les genres les plus impliqués dans ces problématiques. Contrairement aux mycoses, les mycotoxicoses consistent en des intoxications de l'Homme et des animaux via l'ingestion ou l'inhalation de toxines produites par ces moisissures. Historiquement, l'ergotisme a été la première maladie fongique épidémique reconnue en 1676 comme résultant de la consommation d'aliments. Dans les années 1960, la maladie « turkey X disease » qui a décimé des élevages de dindes est l'un des épisodes de mycotoxicoses les plus référencés et signe le point de départ d'une nouvelle discipline « la mycotoxicologie »<sup>1</sup>. D'autres espèces de moisissures sont quant à elles phytopathogènes, entraînant des menaces pour la santé alimentaire, une destruction partielle ou totale des cultures et, par extension, d'importantes pertes économiques. En effet, la quantité de nourriture détruite chaque année en raison d'une infestation fongique est équivalente à la quantité de nourriture suffisante pour 600 millions d'individus, soit 8,5% de la population mondiale<sup>2</sup>.

Le pouvoir pathogène des moisissures résulte en partie de la production par ces microorganismes de métabolites secondaires toxiques appelés mycotoxines. Les métabolites secondaires sont des molécules de structures très variables qui ne sont pas nécessaires à la croissance des moisissures mais qui leur confèrent, en général, un avantage sélectif dans leur environnement. Ces composés ne sont pas nécessairement toxiques et certains sont, au contraire, utilisés en médecine comme antimicrobiens (pénicilline), hypolipémiants (lovastatine), immunosuppresseurs (cyclosporine, acide mycophénolique) ou vasoconstricteurs (ergométrine). En revanche, on recense aujourd'hui plus de 400 mycotoxines. Alors que 30 d'entre elles ont des effets toxiques relativement bien caractérisés pour l'Homme et/ou les animaux, seules six familles de mycotoxines sont réglementées au niveau mondial. C'est le cas notamment de l'aflatoxine qui est l'un des plus puissants agents cancérigènes d'origine naturelle. Les genres *Aspergillus*, *Penicillium* et *Fusarium* sont ceux qui produisent le plus de mycotoxines.

Afin de lutter contre les champignons filamenteux, 21 000 tonnes de fongicides ont été utilisées sur les surfaces agricoles françaises en 2014. Face aux inquiétudes environnementales et sanitaires qui impliquent une diminution de l'utilisation de pesticides d'une part, et au besoin croissant de nourriture nécessaire pour satisfaire une population mondiale estimée à neuf milliards d'individus en 2050 d'autre part, la lutte contre la toxicité de ces agents pathogènes apparaît comme cruciale<sup>3</sup>. La caractérisation des molécules toxiques impliquées dans ces problématiques est une étape préalable à leur contrôle dans les denrées alimentaires. Cependant, on estime que les 11 250 métabolites secondaires fongiques connus aujourd'hui<sup>4</sup> ne représentent que 20% de la totalité des métabolites secondaires sécrétés par les champignons filamenteux.

En Europe et dans les régions tempérées du globe, *P. verrucosum* et *F. graminearum* constituent un risque important pour les productions céréalières. Alors que *P. verrucosum* est un

champignon saprophyte infectant les grains lors de leur stockage, *F. graminearum* affecte les épis de blé lors de leur croissance dans les champs. *P. nordicum* infecte quant à lui principalement les denrées alimentaires protéinées telles que les viandes séchées ou cuites. La caractérisation des métabolites secondaires de ces champignons est essentielle à la diminution des pertes économiques et à la protection des populations. D'autant plus qu'il est suggéré que ces espèces fongiques sont dotées d'importantes capacités métaboliques. En effet, *F. graminearum* est considéré comme l'un des champignons les plus producteurs de métabolites secondaires, et une étude récente a permis de souligner l'importance du potentiel métabolique des champignons du genre *Penicillium* grâce à l'analyse des génomes de 24 de ses espèces<sup>5</sup>.

Au sein du laboratoire, une méthode d'analyse des métabolomes secondaires fongiques a été précédemment développée<sup>6</sup>. Celle-ci repose sur le marquage de l'ensemble des composés produits par les champignons. Dans ce but, les champignons sont cultivés sur des grains de blés marqués afin qu'ils puisent le marquage de ce substrat et l'intègrent dans tous les composés qu'ils génèrent. Cette méthode permet de détecter spécifiquement les molécules d'origine fongique au sein des extraits analysés par spectrométrie de masse et de fournir des informations quant à la composition chimique des molécules analysées.

## 2. OBJECTIFS ET DÉROULEMENT DE LA THÈSE

De nombreuses méthodes existent pour la recherche de nouveaux métabolites secondaires fongiques. Le choix de la méthode d'analyse mise en œuvre au sein d'un tel projet doit être effectué en connaissance du contexte analytique global. Pour cela, une revue des différentes stratégies disponibles pour la recherche de composés fongiques est présentée en introduction.

Ce projet de thèse a été élaboré afin de répondre aux objectifs suivants :

1. Caractériser les métabolomes secondaires de *P. verrucosum*, *P. nordicum* et *F. graminearum*.
2. Appréhender différents outils informatiques disponibles pour assister l'identification de nouveaux produits naturels.
3. Élucider la structure de métabolites secondaires inconnus.
4. Comparer les métabolomes secondaires d'espèces fongiques voisines afin d'apprécier la relation entre proximité phylogénique et similarité métabolique.

Afin de répondre à ces problématiques, ce projet se présente en trois grandes parties distinctes correspondant à l'analyse de chacune des trois espèces fongiques (Figure 1). Dans un premier temps, le métabolome secondaire de *P. verrucosum* est investigué. Grâce à une culture sur grains de blé marqués, il est possible de caractériser les composés produits par ce champignon phytopathogène saprophyte lorsque celui-ci se développe dans des conditions représentatives de son environnement naturel. En effet, puisque le métabolome secondaire des moisissures est finement régulé selon le milieu dans lequel elles se développent, il est important de prendre en considération ce facteur lors de la recherche de composés fongiques. Des outils bioinformatiques permettent de faciliter l'élucidation structurale de composés inconnus parmi les métabolites secondaires détectés. Afin d'apprécier la distribution de la production des nouveaux métabolites secondaires au sein du règne fongique, les composés identifiés peuvent être recherchés au sein des métabolomes secondaires d'espèces voisines et éloignées.

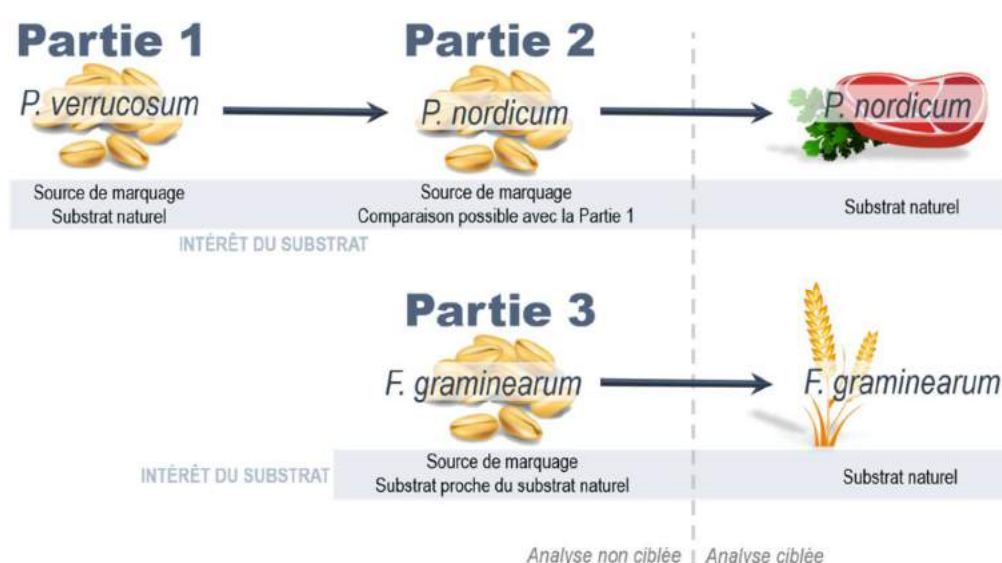


Figure 1. Présentation des travaux expérimentaux réalisés au cours de la thèse.

## Introduction

*P. nordicum* est une espèce phylogéniquement très proche de *P. verrucosum*. Afin de comparer les métabolomes secondaires de ces deux moisissures, les conditions de cultures et d'analyses doivent être identiques. *P. nordicum* est donc cultivé dans un premier temps sur grains de blé marqués. Premièrement, cela permet de rechercher la production par *P. nordicum* des nouveaux composés identifiés chez *P. verrucosum*. Deuxièmement, l'étude des similarités entre les deux métabolomes permet d'apprécier la relation entre proximité phylogénique et similarité métabolique. Enfin, ce substrat modèle constitue une source de marquage nécessaire pour la caractérisation chimique des métabolites secondaires de ce champignon toxigène. Afin de considérer *P. nordicum* dans son contexte environnemental et de prendre en compte l'influence du milieu de culture sur la production de métabolites secondaires, les composés caractérisés grâce au marquage sont ensuite recherchés de manière ciblée lorsque *P. nordicum* se développe sur son substrat naturel.

Dans la troisième partie du travail expérimental mené au sein de cette thèse, le métabolome secondaire de *F. graminearum* est investigué. Puisque ce champignon est phytopathogène, la problématique environnementale dans lequel il s'inscrit est similaire à celle concernant *P. verrucosum*. Dans un premier temps, le champignon est donc cultivé sur grains de blé marqués afin de détecter spécifiquement et de caractériser les molécules qu'il produit sur un substrat végétal. En revanche, contrairement à *P. verrucosum*, *F. graminearum* n'est pas saprophyte et de nouveaux paramètres tels que la réponse du végétal à l'infection fongique sont susceptibles d'intervenir dans sa production de métabolites secondaires. Pour cette raison, les composés caractérisés grâce au marquage sont dans un deuxième temps recherchés dans des échantillons de blé naturellement infestés dans les champs.

### **3. ARTICLE 1 : CARACTÉRISATION DES MÉTABOLITES SECONDAIRES FONGIQUES**

Cette partie présente les différentes méthodes ayant conduit à la découverte de nouveaux métabolites secondaires fongiques. Elle fait l'objet d'une revue soumise pour publication dans *Natural Product Reports*.

Deux types de stratégies sont particulièrement mis en œuvre pour la recherche de nouveaux métabolites secondaires fongiques. D'une part, les méthodes orientées vers l'analyse directe des métabolites secondaires permettent la détection des molécules sécrétées par le champignon dans les conditions expérimentales. D'autre part, l'analyse des génomes fongiques est une stratégie largement utilisée. Des approches transcriptomiques et protéomiques complémentaires peuvent également être mises en œuvre pour appuyer les résultats des analyses métabolomiques et génomiques. Aujourd'hui, le développement rapide de méthodes d'analyses omiques non ciblées permet d'atteindre des niveaux de résolution et de sensibilité qui révolutionnent l'analyse des extraits naturels. Des outils informatiques utilisés en particulier dans l'étude des métabolomes fongiques pour le traitement des importants jeux de données qui en résultent sont présentés dans cette partie. Enfin, l'influence de l'environnement de culture sur la synthèse des métabolites secondaires par les champignons est mise en évidence comme un facteur majeur à prendre en compte pour la recherche de nouveaux métabolites secondaires fongiques. Par la présentation critique de ces outils, cette revue préconise une orientation de la recherche vers une approche intégrée et souligne l'importance du développement des outils bioinformatiques associés.



## From Genomics to Metabolomics, Moving toward an Integrated Strategy for the Discovery of Fungal Secondary Metabolites

T. Hautbergue<sup>a,b</sup>, E. L. Jamin<sup>a,b</sup>, L. Debrauwer<sup>a,b</sup>, O. Puel<sup>a</sup> and I. P. Oswald<sup>a</sup>

Presubmission accepted  
Submitted May 2017

Fungal secondary metabolites are defined by bioactive properties that ensure adaptation of the fungus to its environment. Although some of these natural products are promising sources of new lead compounds especially for the pharmaceutical industry, others pose risks to human and animal health. The identification of secondary metabolites is critical to assessing both the utility and risks of these compounds. Since fungi present biological specificities different from other microorganisms, this review covers the different strategies specifically used in fungal studies to perform this critical identification. Strategies focused on the direct detection of the secondary metabolites are reported since they specifically highlight the expressed part of the fungal genome. Particularly, advances in high-throughput untargeted metabolomics have led to the generation of large datasets whose exploitation and interpretation generally require bioinformatics tools. Then, the genome-based methods used to study the entire fungal metabolic potential are reported. Transcriptomic and proteomic tools used in the discovery of fungal secondary metabolites are presented as links between genomic methods and metabolomic experiments. Finally, the influence of the culture environment on the synthesis of secondary metabolites by fungi is highlighted as a major factor to consider in research on fungal secondary metabolites. Through this review, we seek to emphasize that the discovery of natural products should integrate all of these valuable tools. Attention is also drawn to emerging technologies that will certainly revolutionize fungal research and to the use of computational tools that are necessary but whose results should be interpreted carefully.

### 1. Introduction

Fungi are well known for their capacity to produce a broad diversity of secondary metabolites that provide them with beneficial properties for adequate growth in a fluctuating environment. On the one hand, these properties are of interest to various industries (particularly pharmaceutical, cosmetic and alimentary industries) that either can commercialize the natural compounds directly or develop derived products from the fungal molecules. Newman and Cragg showed that between 1981 and 2006, 28% of the newly developed industrial chemicals were of natural origin and 24% were inspired by natural products<sup>1</sup>. Over 40% of filamentous fungi are presumed to produce antibiotics under natural growth conditions<sup>2</sup>. Since the discovery of penicillin by Alexander Fleming in 1928<sup>3</sup>, medicine has exploited natural fungal defenses against bacteria to protect humans and animals from pathogenic microorganisms. Likewise, many fungal compounds are used as antimicrobials as well as lipid-lowering medications (lovastatin),

immunosuppressants (cyclosporine and mycophenolic acid) and vasoconstrictors (ergometrine).

On the other hand, some fungal secondary metabolites named mycotoxins are the subject of major concern because of their toxicity. Each year, systemic mycoses affecting immunocompromised individuals lead to 1.6-2 million deaths globally<sup>4</sup>. In contrast to mycoses, mycotoxicoses involve intoxication from the exposure to mycotoxins. Ergotism was an epidemic fungal disease recognized in 1676 to result from foods consumption<sup>5</sup>. In the 1960s, the identification of turkey X disease, due to the presence of Aflatoxins represented a turning point in the use of the term mycotoxin<sup>6</sup>, with far more than 400 secondary metabolites being considered as such today<sup>7</sup>. Of these, 30 mycotoxins have been demonstrated to be toxic to humans and/or animals, and only six mycotoxin families are regulated worldwide<sup>8</sup>. Some mycotoxins are also involved in plant diseases as pathogenic or aggressiveness factors, leading to partial or complete destruction of crops and by extension to huge economic losses. The destruction of infected crops was estimated to be equivalent each year to a quantity of food that could feed 600 million individuals, i.e., 8.5% of the world population<sup>9</sup>. Given the environmental concerns seeking to decrease the use of fungicides and the increasing demand for food to feed nine billion people in 2050, the irrepressible development of fungal resistance to pesticides must be faced<sup>4</sup>.

<sup>a</sup> Toxalim, Université de Toulouse, INRA, INP-ENVT, INP-EI-Purpan, Univ. Toulouse 3 Paul Sabatier, 31027 Toulouse, France.

<sup>b</sup> Axiom Platform, MetaToul-MetaboHUB, National Infrastructure for Metabolomics and Fluxomics, 31027 Toulouse, France.



The inherent properties, both beneficial and harmful, of fungal secondary metabolites make the study of these natural products of great importance. Today, 99,000 fungal species are identified and it is estimated that as many as 5 million fungal species exist<sup>10</sup>. Nevertheless, fungal genome investigations of identified species suggest that 80% of their secondary metabolome remains unknown, highlighting the large proportion of compounds waiting to be discovered<sup>11</sup>. Structural elucidation of unknown fungal secondary metabolites is difficult since many of these natural compounds are synthesized in low amounts in very complex matrices. Each secondary metabolite is secreted under particular environmental conditions to adapt the colonization process, resulting in the production of only a small proportion of the total compounds under standard laboratory growth conditions. Moreover, the toxicity of mycotoxins may be observed after long-term exposure or exposure to mixtures with other fungal compounds. For these reasons, mycotoxin characterization should not be restricted to the predominant metabolites and to the observation of acute biological effects.

This review presents an overview of the large spectrum of methods that have been developed to discover fungal natural products (Figure 1) and presents the advantages of each strategy in addressing the abovementioned difficulties. Strategies focused on the direct detection of the secondary metabolites are reported since they specifically highlight the expressed part of the fungal genome. Particularly, advances in high-throughput untargeted metabolomics have led to the generation of large datasets whose exploitation and interpretation generally require bioinformatics tools. Some of the promising algorithms for natural product research are presented herein. Then, the genome-based methods used to study the entire fungal metabolic potential are reported. Transcriptomic and proteomic tools used in the discovery of fungal secondary metabolites are presented as links between genomic methods and metabolomic experiments. Finally, the significant influence of the culture environment on the synthesis of secondary metabolites by the fungus is highlighted. Through this review, we seek to emphasize that the discovery of natural products should integrate all of these valuable tools.

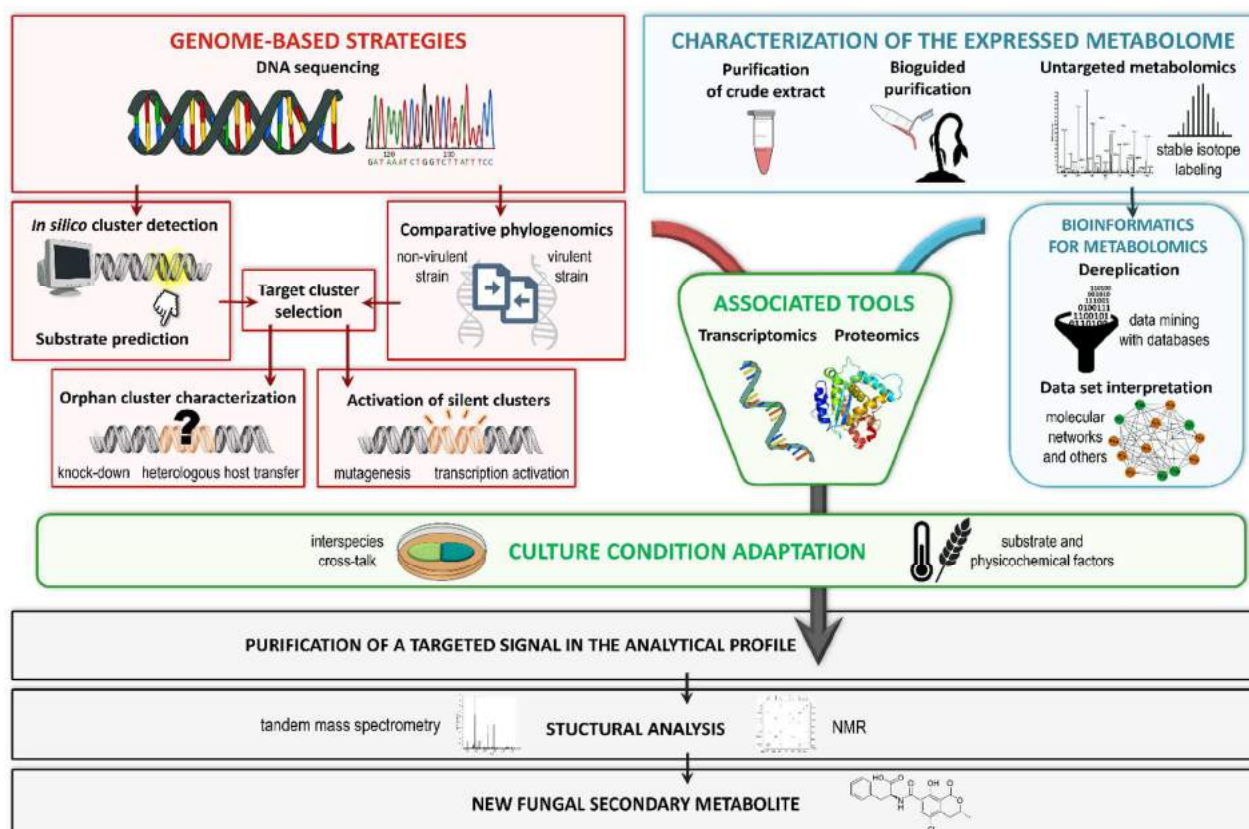


Fig. 1 From genomics to metabolomics, the multiple strategies for the discovery of unknown fungal secondary metabolites.

## 2. Analysis of Expressed Secondary Metabolome

### 2.1 Targeted Purification

**Purification of Crude Extracts.** Many natural products were discovered after observation of an uncharacterized signal in the analytical profile of a fungal extract. Chromatographic purification of the corresponding unknown secondary

metabolite and structural analysis by mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy lead to the characterization of fungal compounds. Among the very numerous examples, two unknown secondary metabolites, arugosins H and G, were isolated from *Emericella nidulans* using normal-phase vacuum liquid chromatography (VLC) and high-performance liquid chromatography (HPLC)<sup>12</sup>. Similarly, two

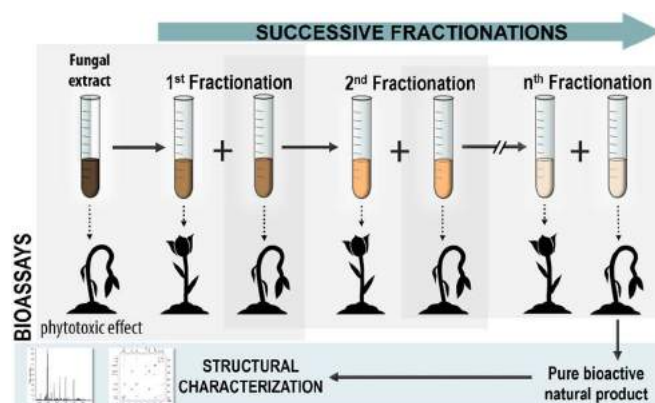
modified peptides were discovered by Garo *et al.*<sup>13</sup>. One of these trichodermamides bears a chlorinated group and displays antimicrobial activities and cytotoxicity against human colon carcinoma. Using this method, Wangun *et al.* identified three unknown antibacterial tetramic acid pyridones from the fruit pathogen fungus *Epicoccum* sp.<sup>14</sup>. Study of the soil fungus *Trichoderma harzianum* revealed the isoharzianic acid, which is responsible for the inhibition of fungal mycelium radial growth, the stimulation of tomato seed germination and the induction of disease resistance<sup>15</sup>. After purification of unknown compounds from an extract of *Aspergillus varicolor*, Wang *et al.* detected 12 varicolorins (A to L)<sup>16</sup>.

The secondary metabolites characterized in these later works displayed biological activities while there was no guarantee about this hypothesis. This strategy is likely unsuccessful in this respect since research on natural compounds is not guided by precise criteria. To guide research on unknown natural compounds, a recent study presented the MeHaloCoA algorithm developed to specifically identify halogenated compounds in untargeted MS data from natural extracts<sup>17</sup>. Many halogenated natural compounds, such as ochratoxin A, a well-known mycotoxin, and the antifungal griseofulvin, display biological effects<sup>18</sup>. A successful application of MeHaloCoA led to the discovery of griseophenone I and chlorogriseofulvin, which have antiproliferative effects<sup>17</sup>. Although this targeting of halogenated products increases the probability of discovering active compounds, it does not guarantee it. Alternatively, the historical strategy for secondary metabolite characterization consists of extractions guided by the observation of biological effects.

**Bioguided Purification.** Several toxic fungal secondary metabolites have been identified during sanitary crises, and their characterization was guided by tracking their toxic effects. Until today, bioguided research on natural compounds has been performed to trace back to the origin of the deleterious effect or of the pharmaceutical activity. Bioguided purification consists of the consecutive fractionation of a fungal extract that retains the studied effect to ultimately obtain the smallest fraction containing only the active substance (Figure 2). In 1988, two fumonisins, B1 and B2, which are the cause of leukoencephalomalacia in horses and hepatocarcinogenic in rats, were discovered from *Fusarium verticillioides* cultures using this approach<sup>19</sup>.

Many examples of fungal natural product identification based on the bioguided strategy have been reported in the literature, such as the discovery of mevinoлин, an HMG-CoA reductase from *A. terreus*<sup>20</sup>, and a tyrosine-derived alkaloid characterized from *F. incarnatum*<sup>21</sup>. Another study on the fruit pathogen fungus *Mollisia benesuada* revealed the synthesis of an antimicrobial, cytotoxic and phytotoxic compound named besudon<sup>22</sup>. In the same way, bioassay-guided fractionation was performed on an extract of *Neosetophoma samarorum* and led to the isolation of four antimicrobial compounds<sup>23</sup>. In an effort to identify anticancer products, Liu *et al.* also used this strategy with *Spicaria elegans* and characterized trichodermamide A, espergillazine A and six cytochalasins<sup>24,25</sup>. The crude extract of a halotolerant fungus displaying cytotoxic effects was fractionated

using liquid chromatography (LC) and led to the discovery of three compounds, varicolorins A to C<sup>26</sup>. Communesins were also characterized in extracts of *Penicillium expansum*<sup>27</sup>. Recently, this strategy was used to test the antioxidant activity of *Bipolaris sorokiniana*, leading to the discovery of sorokinin<sup>28</sup>.



**Fig. 2** Principle of the bioguided fractionation strategy to discover secondary metabolites with biological properties: the fungal extract demonstrating bioactivity (here, a phytotoxic effect) is fractionated, and each resulting fraction is tested for the specific activity. The bioactive fraction is then also fractionated, and each resulting subfraction is tested in turn. This protocol is carried forward until the last subfraction contains the pure bioactive compound that can then be characterized (here, by MS and NMR, for example).

This approach can be extended to the high-throughput screening of thousands of secondary metabolites from many different fungal extracts to identify compounds with a targeted biological effect. For example, Vansteelandt *et al.* selected a *Penicillium* strain after screening for cytotoxicity against cancer cell lines versus non-tumor cell lines. These activities were found to originate from the synthesis of a chlorinated sesquiterpenoid named ligerin<sup>29</sup>. In searching for molecules capable of inhibiting the SecA protein (involved in bacterial resistance to toxins), Parish *et al.* screened extracts of *Geomyces pannorum* and characterized pannomycin<sup>30</sup>. Bioguided purification is suitable for secondary metabolites that have strong activities or are produced in considerable quantities but could be limited by the “cocktail effect” and not be effective in cases where an effect observed in mixed extract is lost in the purified fraction<sup>11</sup>. Nevertheless, this method is commonly integrated in studies of the fungal secondary metabolome as a complement to other strategies reported in this review<sup>31,32</sup>.

## 2.2 Untargeted Metabolomics

Metabolomics consists of the determination of whole metabolites in a biological system. In particular, untargeted metabolomics seeks to obtain a fingerprint of the metabolome of a studied organism using MS or NMR data<sup>33</sup>. Technological advances in MS and NMR methods in recent decades have revolutionized the research on natural compounds by providing high-resolution datasets. At present, this research needs to be directed toward the development of computational algorithms for processing such complex results. Data mining and spectral interpretation remain the critical steps in untargeted research

on secondary metabolites<sup>34,35</sup>, particularly when analyzing complex matrices such as fungal extracts<sup>36</sup>. Analytical strategies, dereplication tools and software developed to assist untargeted studies of new natural products are described hereafter.

**Stable Isotope Labeling.** One strategy to guide the detection of fungal secondary metabolites in complex metabolomic fingerprints is to specifically label fungal products using stable isotope labeling (SIL). When growing on a labeled substrate, fungi produce labeled secondary metabolites that can be distinguished from contaminants and background noise. Depending on the level of specificity of the labeled substrate, the labeling of fungal secondary metabolites can be restricted to a specific family or spread to the entire secondary metabolome. In their experiment, Klitgaard *et al.* studied the secondary metabolites synthesized from phenylalanine by feeding *A. nidulans* with labeled phenylalanine. The resulting extract was compared with the extract of a non-labeled culture. The compounds of interest were specifically detected using MS by exploiting the given mass difference between the secondary metabolites produced via the incorporation of the <sup>13</sup>C<sub>9</sub><sup>15</sup>N-labeled phenylalanine and the secondary metabolites produced from non-labeled amino acids. Several unknown compounds were identified as several analogues of nidulanin A and fungisporin<sup>37</sup>.

In addition to the precursor-focused SIL strategies, labeling the global source of carbon atoms (<sup>13</sup>C) and/or nitrogen atoms (<sup>15</sup>N) leads to the untargeted labeling of all fungal secondary metabolites<sup>38</sup>. As mentioned above for the precursor-focused strategies, the observation of particular isotopic patterns allows the specific detection of fungal secondary metabolites. Moreover, the labeling of fungal metabolites at the atomic scale has the additional advantage of helping to determine the chemical formula of each secondary metabolite. Considering the mass difference between the labeled and unlabeled signals, the number of carbon and/or nitrogen atoms contained in each secondary metabolite can be determined. Bueschl *et al.* cultivated *F. graminearum* on two distinct media, one containing native glucose with almost 100% <sup>12</sup>C and another containing <sup>13</sup>C-glucose as a unique carbon source. Based on the mass difference between the symmetric isotopic patterns of the labeled and non-labeled co-analyzed metabolites, the carbon atom composition of all the fungal compounds was elucidated and used to determine the chemical formulas<sup>39</sup>. Another approach to the <sup>13</sup>C-labeling of fungal secondary metabolites called isotopic ratio outlier analysis (IROA) is based on the preparation of two labeled metabolomes with specific proportions of labeled nutrients from 5% and 95% <sup>13</sup>C-labeled cultures<sup>40</sup>. Some overviews of IROA methods are available<sup>40,41</sup>. Automatic detection of the isotopic patterns can be performed using ClusterFinder for IROA<sup>42,43</sup> and additional data mining is possible with FragExtract<sup>44</sup>. This latter tool was applied in the study of *F. graminearum*, and 9 unknown compounds were identified<sup>44</sup>.

Although calculation of the number of carbon atoms could be sufficient to unambiguously determine the chemical formula of small secondary metabolites (<400 Da), metabolites with

high molecular weights could have several possible formulas, even with measurements of high mass accuracy. To overcome this limitation, Cano *et al.* developed a protocol using double isotope labeling of the fungal substrate with both labeled carbon atoms and labeled nitrogen atoms. This method allows the specific detection of all the fungal compounds as well as the unambiguous determination of their chemical formulas<sup>45</sup>. Three differently labeled wheats were generated after culturing in hermetic chambers alimented with native or <sup>13</sup>C-labeled CO<sub>2</sub> and with native or <sup>15</sup>N-labeled nutrients: (i) a native wheat (containing 99% <sup>12</sup>C), (ii) a <sup>13</sup>C-labeled wheat (containing 97% <sup>13</sup>C) and (iii) a <sup>13</sup>C<sup>15</sup>N-labeled wheat (containing 97% <sup>15</sup>N and 50% <sup>13</sup>C). The harvested wheat grains were then used as only source of carbon and nitrogen atoms for the fungus. After distinct extractions and MS analyses of the three cultures, a list of all secondary metabolites produced by the fungus grown on the wheat grains was generated, and their chemical formulas were unambiguously determined, even for high molecular mass compounds. The second advantage of this method lies in the labeling of the plant with 50% of <sup>13</sup>C and 97% of <sup>15</sup>N, which enables specific and fast detection of each fungal secondary metabolite in the complex mass spectra based on the specific isotopic patterns. This method was successfully validated for the well-documented secondary metabolism of *A. fumigatus*, and a new member of the fumigaclavine family was identified<sup>45</sup>. Similarly, the application of double SIL of a fungal substrate led to the detection of 98 secondary metabolites in *P. verrucosum* and the determination of their chemical formulas. Among them, 82 compounds were unknown<sup>46</sup>.

**Dereplication.** Untargeted metabolomics often leads to complex fingerprints of fungal secondary metabolomes. In this context, dereplication, i.e., the annotation of known compounds in a mixture, is a key step<sup>47,48,49</sup>. Particularly in the field of secondary metabolite research, Bills *et al.* stated that “dereplication is a vital step in natural product discovery processes”<sup>11</sup>. To avoid wasting time rediscovering known metabolites<sup>11</sup>, this early data-mining strategy seeks to highlight known compounds and guide the analysis toward unknowns by submitting the MS or NMR results to databases. Dereplication must be based on the appropriate databases that correspond to the studied organism and are adapted to the type of data obtained (UV, NMR or MS spectra, for example). For fungal studies, databases such as Antibase<sup>50</sup> or the Dictionary of Natural Products<sup>51</sup> list known compounds with the associated metadata such as the chemical formula and exact mass. Since many isomeric products may have the same exact mass, i.e., chemical formula, restricting the databases from global natural products to fungal secondary metabolites could limit the number of matching compounds.

Tandem mass spectrometry (MS/MS or MS<sup>n</sup>) coupled with LC provides both chromatographic retention times and fragmentation pattern specific to each detected secondary metabolite. Hence, some databases reference fungal secondary metabolites with their chromatographic and MS or MS/MS profiles. El-Elimat *et al.* recorded ultra-high performance liquid chromatography (UHPLC) retention times and MS, MS/MS, and UV spectra of 170 known fungal secondary metabolites<sup>52</sup>.

Similarly, Nielsen *et al.* listed the LC-UV-MS properties of 474 fungal secondary metabolites<sup>53</sup>, and Kildgaard *et al.* presented a database of 1300 compounds for the dereplication of marine-derived fungal secondary metabolites<sup>54</sup>. In addition to these homemade databases, the Global Natural Products Social Molecular Networking (GNPS) system is a general tool for the dereplication of natural products. This open-access library shares approximately 220,000 MS/MS spectra representing more than 18,000 natural products from the MassBank<sup>55</sup>, ReSpec<sup>56</sup> and NIST<sup>57</sup> databases as well as reference compounds from approved worldwide contributors<sup>58,59</sup>. Recently, the dereplication algorithm DEREPLICATOR was associated with the GNPS system<sup>60</sup>. This tool, which specializes in the dereplication of peptidic natural products from metabolomic experiments, was created to complete the spectral alignment algorithms for the non-ribosomal peptide (NRP) dereplication of cyclic peptides<sup>61</sup> and the platform for the classical dereplication of NRPs, iSNAP<sup>62</sup>. While such databases increase the dereplication efficiency, full identification is only possible from comparison (of the chromatographic pattern and MS/MS, NMR or UV spectrum, for example) with a standard compound.

In addition to the tens of thousands of MS/MS spectra of known compounds, the GNPS system, as the largest library of publicly shared datasets, containing more than 80 million public, uncharacterized, MS/MS spectra, facilitates unknown compound dereplication. Presented as a data-mediated social network, public datasets deposited by laboratories across the globe can be compared with each other. Thus, unknown secondary metabolites identified in different studies can be linked and knowledge can be shared in the interest of natural product characterization<sup>59</sup>.

**Computational Approaches to Assist Data-Mining in Metabolomics.** Further advances in metabolomic research on natural products has produced large datasets with hundreds of MS/MS spectra per experiment, leading to the development of computational tools that can help in data interpretation. GNPS enables the organization of hundreds of MS/MS spectra according to their similarity, assuming that their similar MS/MS spectra may originate from similar structures<sup>63</sup>. This organization highlights metabolites with potentially similar structures, including knowns and unknowns, and therefore assists in their structural elucidation. Mohimani and Pevzner focused on the application of molecular networks for peptidic natural products. Since these secondary metabolites are primarily produced as group of analogs, GNPS simplifies their detection by clustering their similar MS/MS spectra<sup>64</sup>. Based on latent Dirichlet allocation (LDA), a generative statistical model, the software MS2LDA was also developed to group molecules sharing specific substructures according to their MS<sup>2</sup> profiles<sup>65</sup>. Contrary to the GNPS system, MS2LDA users can specify the neutral losses or mass fragments of interest to avoid grouping secondary metabolites by non-specific fragmentations (i.e., losses of CO and H<sub>2</sub>O).

In addition to these data mining tools, software has been developed to support the interpretation of fragmentation mass spectra and the structural characterization of metabolites. Fragmentation trees (FTs) are data structures generated using

particular computational tools to interpret or compare MS/MS spectra<sup>66,67</sup>. Each node in an FT represents an ion (precursor or fragment), and the edges illustrate the fragmentation reactions. Such an algorithm was first designed by Rasche *et al.* based on known fragmentation rules<sup>68</sup> but was limited by the tremendous number of rules and the fact that totally unknown compounds could fragment according to unknown rules. To address this limitation, Rasche *et al.* created an FT alignment algorithm that compares standard FTs with uncharacterized FTs to highlight the structural similarities between known and unknown natural products<sup>69</sup>. Pairwise alignment similarity matrices of standard FTs showed good correlation with the structural similarities, validating this method. Therefore, this completely automated and “rule-free” analysis is considered a guide to highlight the structural similarities between standards and unknown secondary metabolites. The main limitations of FTs lie in the generation of MS<sup>n</sup> experiments from low abundances of parent ions, particularly in the context of natural product analyses. Caution must also be used because the fragmentation pattern of a molecule depends on the mass spectrometer and fragmentation parameters used. While the GNPS software compares the fragmentation patterns generated from a single analysis or from analyses performed using the same mass spectrometer, FT alignment compares unknown FTs with standard FTs from databases that were generated using different equipment. Users should ensure to not compare the incomparable.

In terms of the structural elucidation of peptidic natural products, classical proteomic tools for peptide sequencing can fail due to the incorporation of non-proteinogenic amino acids and the recurring important structural modifications. To overcome these limitations, NRPquest was created to compare the MS/MS spectra of multiple peptidic natural products and was successfully used for the *de novo* sequencing of peptides<sup>70</sup>. In addition, the freely available open-source software mMass includes a specific algorithm for the *in silico* fragmentation of peptidic natural products to assist the interpretation of mass spectra of linear and cyclic peptides<sup>71</sup>.

Some other bioinformatics tools, such as MassTRIX<sup>72</sup>, CSI:FingerID<sup>73</sup>, CFM-ID<sup>74,75</sup>, MAGMa<sup>76</sup>, MetFrag<sup>77</sup> and MassFrontier, have been developed to support the structural elucidation of unknown compounds, but not specifically natural products. Although these tools may provide valuable assistance, predictions from all these computational tools need to be carefully considered.

### 3. Proteomics

In the quest of fungal product discovery, proteomic approaches can be implemented for direct analysis of the enzymes involved in the synthesis of the secondary metabolites. The specificity of such analyses results from the fact that fungal secondary metabolites are mainly synthesized by core enzymes belonging to well-defined families, such as non-ribosomal peptide synthases (NRPSs), polyketide synthases (PKS) and dimethylallyltryptophan synthases (DMATs)<sup>78–80</sup>. Although these proteomic approaches are relevant for the exploration of enzymes involved in the secondary metabolome, they are more

a way to support one or several of the methods mentioned in this review, rather than fully fledged strategies for the discovery of fungal secondary metabolites.

### 3.1 Gel-Based Proteomics

Using comparative 2D polyacrylamide gel electrophoresis (PAGE)-LC-MS/MS analyses, Owens *et al.* studied the proteome variations of *A. fumigatus* under different conditions of oxidative stress. The proteins in the crude extract were separated by 2D sodium dodecyl sulfate (SDS)-PAGE. Proteins with quantitative differences between the experiments were excised, digested by trypsin and analyzed by LC-MS/MS. The results demonstrated the presence of approximately 20 enzymes that were involved in the secondary metabolome and specifically increased during oxidative stress<sup>81</sup>. A similar approach, called the Proteomic Investigation of Secondary Metabolism (PrISM), targets the core enzymes of the secondary metabolome by selecting the high-molecular weight proteins after 2D SDS-PAGE separation. After a tryptic digest, NRPSs and PKSs are specifically detected in MS/MS data by tracking the neutral losses corresponding to the labile posttranslational modifications of the specific domains of these enzymes. After characterization of the enzymes, their PCR (polymerase chain reaction) primers are designed, the corresponding gene clusters are identified, and the corresponding secondary metabolites are characterized<sup>82,83</sup>. Of note, gel-based proteomics methods are often restricted to the analysis of abundant proteins<sup>84</sup>.

### 3.2 Non-Gel-Based Proteomics

To study proteins of low abundance, a strategy involving the use of isobaric tags for relative and absolute quantification (iTRAQ) was developed<sup>85</sup>. Concerning fungi, this technique has the advantage of enabling the analysis of proteins secreted in low amounts (for example in the early stage of plant infection) and the quantitative comparison of 4 samples in the same analysis<sup>86</sup> (to compare different culture conditions or different growth stages, for example). This strategy was used by Taylor *et al.* to detect the biosynthetic pathways specifically upregulated by *F. graminearum* when growing in conditions favorable to mycotoxin production<sup>86</sup>. Labeling digested proteins followed by LC-MS/MS analysis led to the observation of 18 proteins strongly upregulated during the mycotoxin production stage. This method to discover numerous candidate pathogenicity-related proteins is an interesting starting point for the characterization of pathogenic fungal products. Other methods have been developed to improve the interpretation of complex mass spectra generated from protein extracts but have not been applied in fungal studies. Among others, the orthogonal active site identification system (OASIS) and the activity based protein profiling (ABPP) consist of the derivatization of the active site of NRPSs and PKSs with a specific chemical probe, followed by an affinity purification to eliminate the MS signals of all proteins not involved in secondary metabolism<sup>87,88</sup>.

## 4. Genome Mining for Fungal Product Discovery

Thanks to technological advances in molecular biology, the sequencing of microbial genomes has substantially increased in recent decades. Observation of these fungal genomes has led to the establishment of two general principles: (i) genes involved in the synthesis of secondary metabolites are mostly organized in clusters that are principally located in the non-syntenic regions of subtelomeric extremities of the genomes<sup>89</sup> and (ii) these genes code for core enzymes belonging to well-defined families such as NRPS, PKS and DMATS<sup>78–80</sup>. The observation of orphan clusters, defined by Gross *et al.* as gene clusters whose corresponding secondary metabolites are still undiscovered<sup>90</sup>, indicates that more than 80% of the fungal compounds remain to be identified. This unexplored portion is partly explained by the fact that many gene clusters, called silent clusters<sup>90</sup>, are not expressed under standard laboratory culture conditions. The misemployment of the term “cryptic clusters” to characterize silent clusters was exposed by Harald Gross who states that “gene clusters are more of cryptic nature due to their undiscovered characters, but not in a sense of silence”<sup>90</sup>. To avoid confusion and ensure the accuracy of the information, only the terms orphan and silent will be used in this review to specify the status of unknown clusters. Based on genetic observations, a strategy for fungal secondary metabolite discovery consists of the targeted analysis of orphan clusters and more specifically of the awakening of silent clusters. Whereas the abovementioned studies directly targeted the secondary metabolites by specifically focusing on the expressed part of the genome, genetic engineering allows the characterization of every unknown fungal gene cluster regardless of whether it is expressed in the natural environment. Several reviews report the different genome mining tools employed in natural products research<sup>91–95</sup>. In this section, all methods employed in fungal genome studies are reported and discussed. First, several algorithms developed to detect the gene clusters involved in the secondary metabolome are described. Then, knock-out experiments to characterize the natural products from orphan clusters are detailed. Finally, we focus on strategies elaborated to activate the transcription of silent gene clusters.

### 4.1 Gene Cluster Analysis

***In silico* Cluster Detection.** The first step of genome mining strategies involves the detection of gene clusters involved in the synthesis of secondary metabolites. The mostly rule-based bioinformatics algorithms developed for the untargeted detection of these clusters from sequenced fungal genomes are described in some specific reviews<sup>96,97,98</sup>. Owing to the identification of genes coding for well-conserved protein domains related to the synthesis of secondary metabolites, these approaches are very effective at detecting gene clusters that code for known biosynthetic pathways with high precision<sup>96</sup>. SMURF identifies putative clusters by recognizing the conserved sequences of the three major core enzymes, NRPSs, PKSs and DMATSs and then evaluates their adjacent genes to test their inclusion in the cluster<sup>99</sup>. Similarly, ClustScan searches all the PKS, NRPS or hybrid PKS-NRPS sequences in a

semi-automatic manner<sup>100</sup>. AntiSMASH was developed based on the same principle but has the advantage of analyzing 44 known classes of biosynthesis gene clusters of secondary metabolites<sup>101,102,103</sup>. Recently, PRISM was developed to identify type II PKSs and a wide range of enzymes involved in tailoring reactions<sup>104</sup>. Finally, Wolf *et al.* created a tool called SMIPS-CASSIS to take into account the possible co-regulation of clusters<sup>105</sup>. Based on the hypothesis that some co-regulated clusters include promoters sharing the same transcription factor binding sites, the cluster prediction is guided by the investigation of areas around the core genes, where the promoters share a common binding site.

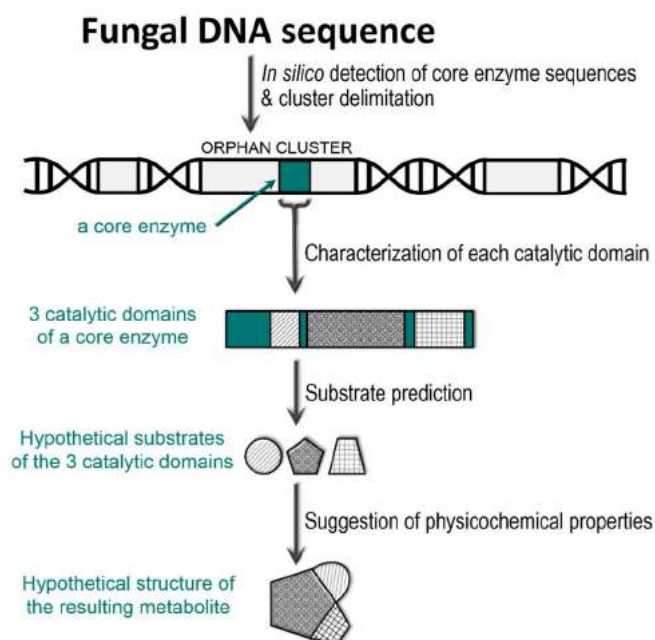
To identify clusters without core enzymes, some motif-independent protocols have been developed. Takeda *et al.* created an algorithm called MIPS-CG that is based on comparison of the DNA sequences of two genomes<sup>106</sup>. This tool detects the clusters involved in the secondary metabolome by searching for a similar order of genes and considers their high enrichment in non-syntenic blocks. MIDASS-M integrates genomic and transcriptomic data to highlight the cooperatively regulated genes that are located in the same region of a genome and led, for example, to the discovery of a known cluster that was three time longer than that previously predicted by SMURF<sup>107</sup>. These two methods of motif-independent cluster detection are reviewed in detail by Umemura *et al.*<sup>108</sup>.

To conclude, it is important to bear in mind that these bioinformatics tools are predictive and that every cluster hypothesis should be confirmed experimentally. These algorithms are very useful for quickly acquiring a general view of an uncharacterized fungal genome but could miss important information such as genes intertwined in superclusters<sup>109</sup>.

**Comparative Phylogenomics.** Phylogenomics couples of genomics and phylogenetics to predict the role of a gene. This method is based on the theory of timeline evolution and focuses on the distribution of genes on a phylogenetic tree<sup>110</sup>. Comparative phylogenomics has been applied to detect gene clusters involved in the synthesis of bioactive secondary metabolites. In an effort to identify genes likely involved in the phytopathogenicity of *Stagonospora nodorum*, Syme *et al.* compared the genomes of three strains: a virulent natural strain, an avirulent strain, and a pathogenic strain that exhibited virulence not observed in the natural strain. This intraspecific genome comparison led to the identification of genes present in the virulent strains but absent in the avirulent strain<sup>111</sup>. Similarly, a phylogenomic comparison of strains of a barley pathogen, *Cochliobolus* spp., revealed clusters specific to particular pathotypes. Deletion of these clusters confirmed their involvement in the virulence<sup>112</sup>. This comparative genomics strategy to highlight clusters of interest was also used in interspecific experiments to study the horizontal transfer of clusters related to virulence. Through comparative phylogenomics, Bushley *et al.* noted modular rearrangements and recombinations in *C. heterostrophus* that are sustained by the NRPS involved in the AM-toxin synthetase of *Alternaria alternata*. These modifications could impact the natural product

structure, leading to unknown compounds with enhanced or reduced toxic properties<sup>113</sup>.

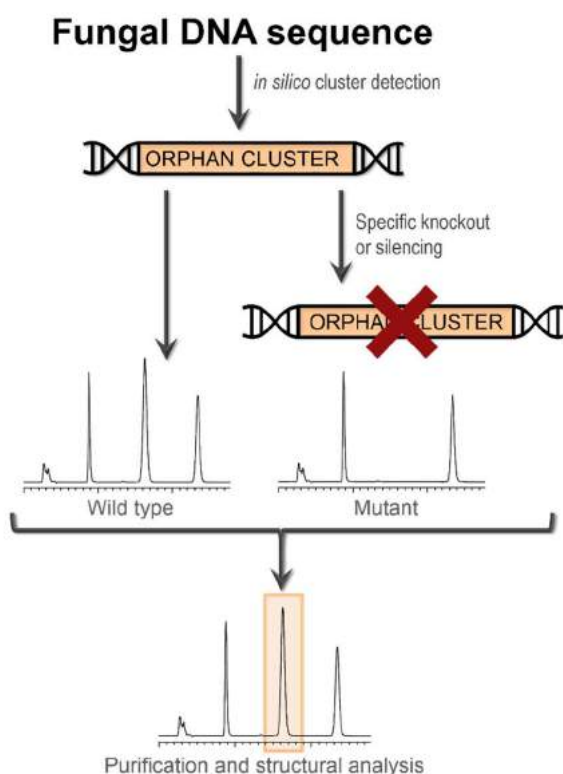
**In silico Prediction of Cluster Products.** NRPSs and PKSs include several domains (A-domains for NRPSs and AT-domains for PKSs) involved in the specific selection of the building blocks incorporated during the synthesis of the secondary metabolites<sup>114</sup>. Several algorithms have been developed to detect and identify all the substrate-binding sites in a cluster. These bioinformatics tools may suggest the structural features of the resulting secondary metabolites by predicting the different precursor units involved in their synthesis (Figure 3). Among the substrate predicting tools, the ASMPKS database proposes a polyketide prediction from unknown PKSs<sup>115</sup>. Moreover, the program NP.searcher can consider the tailoring reactions<sup>116</sup>. Other bioinformatics servers reviewed by Anand *et al.*<sup>117</sup> are available, such as SEARCHPKS<sup>118</sup>, NRPS-PKS<sup>119</sup>, SEARCHGTr<sup>120</sup>, PKS/NRPS Analysis<sup>121</sup>, SBSPKS<sup>122</sup>, NRPSpredictor2<sup>123</sup> and NRPSsp<sup>124,125</sup>. By increasing the number of reactions and potential substrates considered, the open-source application PRISM improves the accuracy of the structure prediction<sup>104</sup>. A similar approach developed by Minowa *et al.* also includes a program for suggesting the molecular activity<sup>126</sup>. While these tools are able to characterize the different catalytic sites of proteins from their genes, one of their major limitations lies in the prediction of the iteration, that is to say the skipping of catalytic steps during the synthesis of a metabolite by a single enzyme, particularly for PKSs.



**Fig. 3** Illustration of the bioinformatics strategies to predict NRPS- or PKS substrate specificities and to suggest the structure of the resulting secondary metabolite. After in silico detection of the sequence of core enzymes and the delimitation of the orphan cluster, characterization of the substrate-binding sites of each core enzyme (A-domain for NRPS and AT-domain for PKS) allows the assumption of all building blocks incorporated into the secondary metabolites and, consequently, their physicochemical properties

## 4.2 Knock-out of Target Genes

**Linking Genes and Secondary Metabolites.** Once a cluster of interest is detected, the characterization of its products involves stopping or reducing its expression. A signal that is present in the analytic profile of the wild strain but absent or reduced in the analytic profile of the repressed strain can be attributed to the corresponding secondary metabolite. After locating the target signal, purification and structural characterization of the molecule must be performed for the expressing strain (Figure 4).



**Fig. 4** Specific repression of an orphan cluster and comparison with the wild-type fungus facilitates locating the corresponding secondary metabolites in the analytical profile. Then, purification and structural analysis steps are performed.

Knock-out organisms can be created by deletion of the studied sequence via homologous recombination. For example, the mutant resulting from a double knock-out in *A. nidulans* led to the discovery of four emericellamides C-F<sup>127</sup>. However, this process is tedious for fungal genomes since fungi are known for a low frequency of well-processed homologous recombination. The incorrect replacement of a target gene by the suitable gene cassette can be partly explained by the non-homologous end-joining pathway that frequently occurs in fungi<sup>128,129</sup>. Inhibition of the Ku70/Ku80 dimeric protein complex involved in repairing DNA double-strand breaks drastically increases the gene targeting efficiency<sup>130,131</sup>. Some recent technologies based on the recruitment of bacterial nucleases were also developed to overcome this limitation. To the best of our knowledge, no article has reported the discovery of fungal secondary

metabolites using these methods, but the CRISPR (clustered regularly interspaced short palindromic repeats)-Cas9 system stands out as a promising tool, as discussed in the conclusion of this review.

**Deleting Highly Expressed Genes to Decrease Background Noise.** Multiple reasons may explain why some clusters remain uncharacterized. The difficulty of purifying and structurally analyzing some natural products may be a limiting factor, and the production of a major secondary metabolite by a fungus may hide the expression of less abundant metabolites. In some studies, several highly expressed genes were therefore deleted from the fungal genome to decrease the background noise and favor the detection of other compounds. Chiang *et al.* deleted eight of the highly expressed genes in *A. nidulans* and characterized aspercryptin<sup>132</sup>. Similarly, asperfuranone was discovered after deletion of the highly expressed sterigmatocystin gene in *A. nidulans*<sup>133</sup>.

## 4.3 Heterologous Host Transfer

The transfer of orphan clusters in non-natural hosts represents another characterization strategy for cluster products. By cloning an entire gene cluster from a fungus, the host may express the entire secondary metabolite pathway. The difference observed between the wild host and the mutant can therefore be attributed to the expression of the transferred DNA sequence. The ease of establishing a link between a cluster and a natural product once the sequence is expressed by the host is a great advantage<sup>134</sup>. Heterologous host transfer is particularly suitable when the studied fungus is difficult to grow under laboratory conditions or cannot be genetically manipulated<sup>135</sup>. Furthermore, the host genome can be deleted of all of his own clusters before acquiring the metabolite pathway of interest to eliminate interferences and facilitate metabolite identification<sup>79,136</sup>.

*A. nidulans* was presented in several studies as an appropriate host since its metabolome is well characterized<sup>137</sup>. This fungus produces only a few secondary metabolites, providing a stable and low background for heterologous expression<sup>136,138</sup>. After *in silico* analyses of dermatophyte fungal genomes, Yin *et al.* observed a cluster of genes similar to that involved in the production of the immunosuppressant neosartoricin. To investigate the production of metabolites with potential therapeutic interest, the researchers transferred the cluster to *A. nidulans* and discovered neosartoricin B<sup>139</sup>. In another study, a gene similar to a 6-methylsalicylic acid (6-MSA) synthase was observed in the genome of *A. aculeatus*. Heterologous expression in *A. nidulans* combined with substrate labeling showed that <sup>13</sup>C-labeled 6-MSA was incorporated to produce three unknown secondary metabolites named aculins A and B and epi-aculin A<sup>140</sup>. The use of bacterial hosts is another possible strategy for the heterologous expression of fungal genes but presents significant limitations. First, the removal of introns is required but leads to weaker metabolite production<sup>135</sup>. Second, some bacterial hosts such as *Escherichia coli* cannot activate heterologous PKSs and NRPSs

using their own intrinsic phosphopantetheinyl (Ppant) transferases<sup>141</sup>. Finally, the broad chemical diversity of fungal PKSs and the consequent requirement of a wide variety of substrates to generate fungal secondary metabolites necessitates the transfer of all the enzymatic machinery involved in the synthesis of such building blocks from the fungus to the bacterial host<sup>141</sup>.

Even if gene transfer to heterologous hosts offers clear benefits and is widely used in bacterial studies, implementing these methods could be very laborious for fungal genomes<sup>142</sup>. Significant limitations remain, due to both the large size of some clusters (>40 kb), making them impossible to transfer in a single plasmid or BAC vector<sup>143</sup>, and the presence of potential cross-chemistry between clusters on different chromosomes. Moreover, gene transfer often requires exchange between the native promoter and an inducible host promoter<sup>134</sup>. To overcome some of these limitations, Chang *et al.* described a tool that enables the integration of an unlimited number of genes in *A. nidulans*<sup>136</sup>. At each insertion of a gene, a targeted marker is removed and used for the subsequent insertion. Moreover, Bok *et al.* recently developed the first fungal artificial chromosome (FAC) system for the transfer of entire clusters from *A. terreus* to *E. coli* and *A. nidulans*. This process allows the heterologous expression of large DNA fragments (>150 kb) and led to the discovery of tereazine D<sup>144</sup>.

#### 4.4 Awakening of Silent Gene Clusters by Stimulating Transcription

The knock-out and heterologous host transfer strategies were reported as characterization methods for orphan clusters. In addition, several protocols that focus more specifically on the characterization of silent clusters are intended to awaken silent gene clusters by stimulating their transcription.

**Specific Pathway Overexpression.** Several mechanisms may regulate gene expression in a specific or global pathway. The advantage of the specific modulation of the expression of a gene lies in the fact that it allows the production of the targeted secondary metabolite in significant quantities (for the detection and the structural elucidation) and the establishment of a strong link between the cluster and the natural product<sup>145</sup>. Since a single transcription factor controls the expression of all genes together in a cluster<sup>146</sup>, the overexpression of only one element leads to the transcriptional modulation of the entire target cluster, simplifying cumbersome genetic manipulations.

- **Stimulation of Specific Transcription Factor Activity.** Many strategies have been developed to stimulate the activity of specific transcription factors. First, some studies multiplied the number of gene copies of the specific transcription factors (Figure 5a1). The efficiency of this method was demonstrated in *A. parasiticus* when Chang *et al.* showed that the inhibited synthesis of aflatoxin can be restored by adding copies of the *afIR* gene, which codes for the specific transcription factor AfIR<sup>147</sup>. Second, specific transcription factors can be overexpressed by coupling their genes with an inducible promoter (Figure 5a2)<sup>148</sup>. The inducible alcohol dehydrogenase

promoter (*alcAp*) is mainly used, but a highly expressed constitutive promoter could also be considered. This strategy requires the integration of a small sequence into the fungal genome, which bypasses the limitation of homologous recombination<sup>142</sup>.

Aspyridones A and B were characterized in *A. nidulans* by coupling the *apdR* gene to the inducible promoter *alcAp*<sup>149,133</sup>. Likewise, three novel fusarielins were characterized in *F. graminearum* by replacing the promoter of the specific transcription factor with the highly expressed constitutive promoter *gpdA*<sup>150</sup>. This latter promoter was also used to activate the expression of a specific regulator in *A. fumigatus* and *Neosartorya fischeri*, leading to the characterization of neosartoricin<sup>151</sup>. Finally, apicidin F, an antimalarial compound, was identified in *F. fujikuroi* after overexpression of the specific regulator ASP2 with the promoter of the  $\beta$ -tubulin gene of *F. graminearum*<sup>152</sup>. This strategy was combined with heterologous host transfer and deletion of highly expressed genes by Chiang *et al.* This led to the characterization of asperfuranone for the first time in this fungal species<sup>136</sup>.

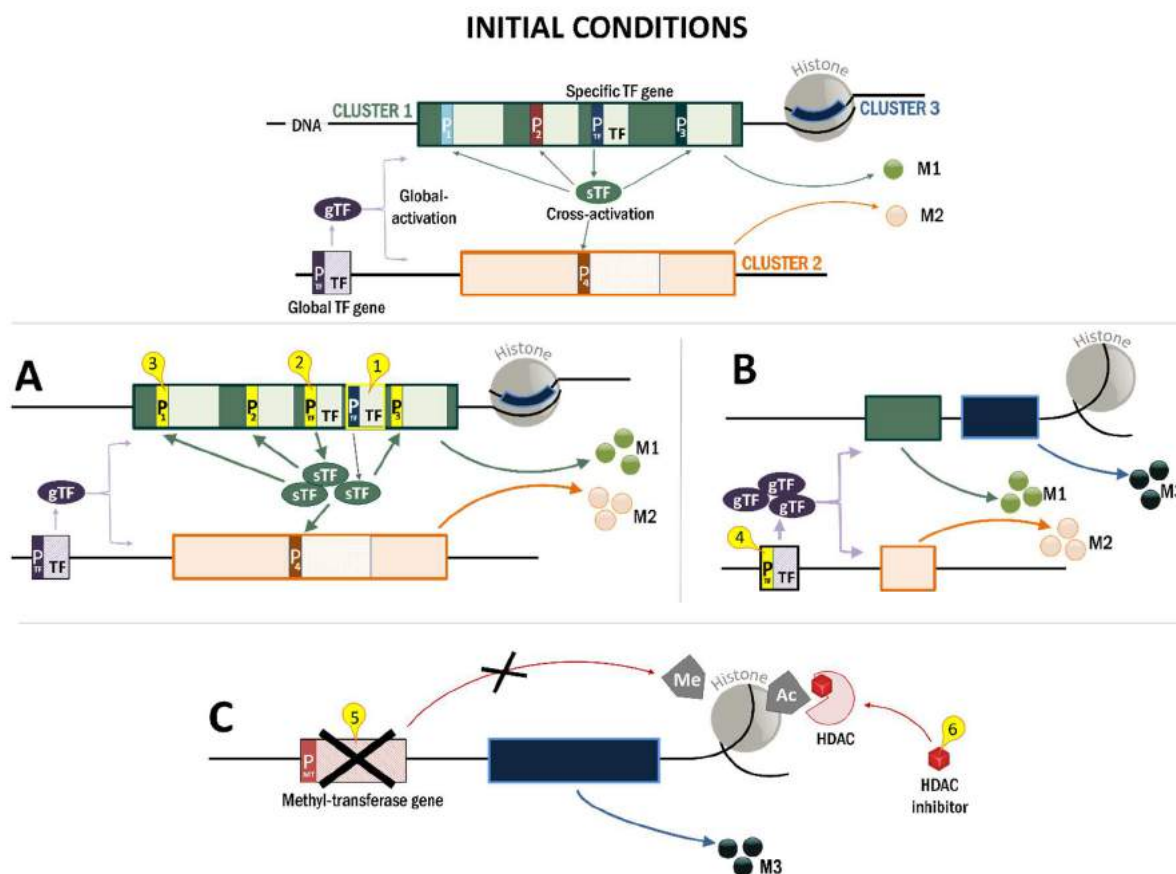
- **Promoter Exchange for the Core Enzyme Genes.** Some transcription factors are regulated posttranscriptionally and many clusters involved in the fungal secondary metabolism are not under the control of easily identifiable transcription factors, which is why some studies have focused directly on the overexpression of the promoters of genes coding for the different core enzymes (Figure 5a3). In fungal genomes, this strategy can be laborious since the replacement of the promoter of all genes is sometimes necessary to stimulate the expression of a cluster<sup>149</sup>. However, in some cases, a single gene determines the expression level of the cluster. In searching for all the silent cluster products in *A. nidulans*, Ahuja *et al.* attempted to replace the promoters of all the transcription factors according to the abovementioned method, but the novel compounds were insufficiently produced for characterization<sup>153</sup>. Next, they exchanged the eight uncharacterized NR-PKS promoters with the inducible *alcA* promoter, resulting in the discovery of seven secondary metabolites with structures derived from  $\beta$ -resorcylaldehyde<sup>153</sup>. Similarly, Yeh *et al.* highlighted the production of microperforanone by *A. nidulans* by exchanging the native promoter of all 13 silent NRPS-like genes<sup>154</sup>.

**Global Transcriptional Modulation.** In the absence of a specific regulation, silent clusters are awakened by global stimulation of the transcription. Global transcription factors are coded by genes in non-cluster areas and also regulate genes not involved in secondary metabolism (Figure 5b)<sup>146</sup>. The most referenced global regulator in terms of fungal secondary metabolism is LaeA<sup>155</sup>. *A. fumigatus* mutant experiments showed that this factor is involved in the transcriptional stimulation of 40% of the biosynthesis genes of fungal secondary metabolites<sup>156</sup>. Wiemann *et al.* even argued that there is no other known global regulation factor that awakens as many silent clusters as LaeA and its homologous do<sup>79</sup>. By modulating of LaeA in *A. nidulans*, Bok *et al.* characterized terrequinone A, an antitumor compound never identified previously in this



fungus<sup>157</sup>. VeA is another global regulator specific to fungi that is involved in secondary metabolite regulation and could also be manipulated to explore natural products<sup>158,159</sup>. Despite the significant use of global transcriptional modulation for secondary metabolite discovery, the mode of action by which these global transcription factors positively or negatively

control metabolite synthesis remains uncertain. Since LaeA was suspected to possess methionine methylation<sup>160</sup> and histone acetylation functions<sup>161</sup>, it is suggested that the transcriptional regulation of cluster expression is regulated by LaeA through epigenetic modulations.



**Fig. 5** Strategies for the transcriptional modulation of fungal clusters. The modifications are expressed in comparison with the initial conditions (scheme on top). **A.** Specific transcriptional modulation: Cluster 1 can be specifically overexpressed by (1) multiplication of the number of gene copies of the specific transcription factor (sTF), (2) by the replacement of the promoter ( $P_{TF}$ ) of the specific TF with an inducible promoter or a highly expressed promoter, or (3) by the replacement of the promoters of all the genes involved in the metabolite synthesis ( $P_1$ ,  $P_2$  and  $P_3$ ). Cross-activation may lead to the activation of another cluster (here, cluster 2) by a specific TF. **B.** Global transcriptional regulation: overexpression of a global transcription factor (gTF) by replacement of the promoter ( $P_{TF}$ ) induces a global stimulation of secondary metabolites from distinct clusters (M1, M2 and M3) (4). **C.** Epigenetic transcriptional modulation: The expression of silent clusters can be stimulated by (5) the inhibition of histone methylation (Me) via the inhibition of DNA methyl-transferase or (6) by the stimulation of histone acetylation (Ac) via the inhibition of histone deacetylase (HDAC).

**Mutagenesis Strategies.** Mutagenesis is another strategy employed to stimulate the expression of silent clusters that lead to the discovery of fungal secondary metabolites. The most widely used mutagenesis strategy for gene cluster awakening consists of inducing random mutations by UV irradiation. One of the first fungal experiments was conducted with *Sphaeropsidales* sp. (strain F-24'707), and UV irradiation led to 25 different mutants. Comparison of the metabolic profiles showed that one of the mutants produced a novel metabolite, mutolide<sup>162</sup>. In addition to UV irradiation, ribosome engineering is another mutagenesis strategy. Before being applied to fungal studies, this principle for the activation of silent genes was described in bacteria<sup>95</sup>. The induction of antibiotic resistance, acquired by mutations of ribosomes (gentamycin resistance) or

mutation of RNA polymerase (rifampicin resistance), was correlated to an awakening of silent gene clusters involved in the secondary metabolome<sup>163</sup>. As fungi are not naturally sensitive to antibiotics, antibiotic resistance has been addressed in several different ways. Drug resistance linked to the awakening of silent clusters was induced by incubation of the fungus with diethyl sulfate (DES) to characterize the secondary metabolome of *P. purpurogenum*. Nine antitumor compounds and three rare chromones were discovered<sup>32,164,165,166</sup>. In another approach, dimethyl sulfoxide (DMSO) was combined with antibiotics to increase the fungal membrane permeability to these molecules and allow the development of drug resistance in *P. purpurogenum*. The incubation of fungal spores with DMSO and gentamicin or neomycin created mutants with

anti tumor activity and led to the characterization of five secondary metabolites, including the recently identified penicimutide<sup>167,168,169</sup>. Finally, membrane permeability and neomycin-resistance were induced in *A. versicolor* using an ultrasound-mediated approach, and the induced aminoglycosides resistance was linked to the synthesis of unknown antitumor compounds<sup>170</sup>. Mutation mediated by antibiotic resistance was performed by Jin *et al.* with a co-culture of *F. semitectum* with *Agrobacterium tumefaciens* in the presence of acetosyringone, a compound known to enhance the bacterium-mediated gene transformation of organisms. Among the resulting mutant fungi, hygromycin B-resistant transformants were highlighted, and apicidin E was characterized<sup>171</sup>.

Mutagenesis strategies are based on the hypothesis that random mutations may awaken silent clusters, but the location of the mutation in the DNA sequence has not been described. It has not been demonstrated that the compounds result only from the transcriptional activation of a cluster. Random mutation may interrupt a biosynthetic pathway, highlighting some of the precursors<sup>172</sup>. Certain drug resistance could be explained by the mutation of enzymes that detoxify the antibiotics<sup>173</sup>. Such enzymes could therefore modify known secondary metabolites to generate new compounds.

#### 4.5 Awakening of Silent Gene Clusters by Epigenetic Engineering

The epigenetic modulation of fungal genomes has been considered an efficient way to control gene expression since it was observed in *A. nidulans* that the deletion of *hdaA*, which codes for an HDAC, generates strong transcriptional activation of sterigmatocystin and penicillin genes<sup>174,175</sup>. Two types of strategies can be used to control the epigenetic factors. First, the use of chemical inhibitors has the advantage of not requiring genetic manipulation (Figure 5c6)<sup>176</sup>. This strategy, called chemical epigenetics, consists of using different types of chemical chromatin modifiers to modulate histone acetylation, methylation, phosphorylation, ubiquitination, sumoylation, glycosylation, or neddylation<sup>80,134,148</sup>. Second, several genes involved in histone modulation can be overexpressed or repressed (Figure 5c5). Among the different epigenetic processes, the two that are primarily modulated are histone methylation (i.e., repression of transcription) and histone acetylation (i.e., stimulation of transcription).

**Modification of Histone Methylation.** 5-Azacytidin is a DNA methyltransferase (DNMT) inhibitor commonly used in research on secondary metabolites<sup>134</sup>. The first antidiabetic and anti-inflammatory compound produced by a microorganism, (S)-(+)-sydonol, as well as three analogous compounds were discovered in *A. sydowii* after the addition of this inhibitor<sup>177</sup>. Culturing fungi with this DNMT inhibitor enabled the characterization of oxylipins produced *de novo* in *Cladosporium cladosporioides*; the identification of lunalides A and B in

*Diatrype* sp.<sup>178</sup>; and the discovery of two secondary metabolites in *P. citreonigrum*, atlantinones A and B<sup>179</sup>.

In addition to these epigenetic modifier additions in the culture medium, mutants have been created by deletion of natural chromatin modifier genes<sup>180</sup>. Recently, knocking out KMT6, a histone methyltransferase in *F. fujikuroi*, activated the expression of four silent genes coding for core enzymes. After heterologous recombination of the uncharacterized sequence in *E. coli*, a novel sesquiterpene was identified<sup>181</sup>.

**Modification of Histone Acetylation.** The synthesis of secondary metabolites can be stimulated by activation of histone acetylation (Figure 5c6). Use of the HDAC inhibitor suberoylanilide hydroxamic acid (SAHA) led to the discovery of cladochromes F and G from *C. cladosporioides*<sup>178</sup> and nygeron A from *A. niger*<sup>182</sup>. Testing different ways to maintain histone acetylation, Albright *et al.* cultivated *A. nidulans* in the presence of diverse HDAC inhibitors (SAHA, trichostatin A or sodium butyrate) and identified a fellutamide and two antibiotics never characterized previously in this fungus<sup>183</sup>.

Another approach to controlling protein acetylation under laboratory conditions involves genetic engineering. In the above study, Albright *et al.* transferred to the fungus an inducible acetylase that remains silent in the absence of stimulation. The secondary metabolomes of the induced and repressed genomes were compared, and the results confirmed the results from the experiments with the HDAC inhibitors.

## 5. Transcriptomic Approaches

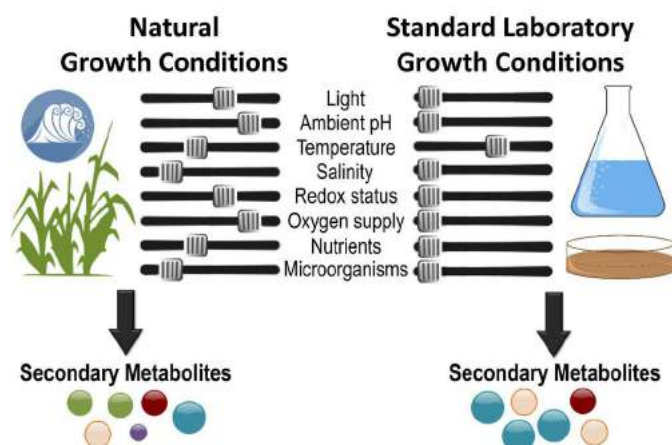
Fungal secondary metabolites can be discovered by genomic engineering of orphan clusters or by direct observation of the fungal secondary metabolome, but in several studies, the transcriptome is an intermediate that helps in understanding the links between the genome and metabolome. In the proteomic approaches, the different transcriptomic techniques are a way to support one or several of the abovementioned methods rather than fully fledged strategies for the discovery of fungal secondary metabolites. In this context, microarray and RNA-seq (or whole transcriptome shotgun sequencing) methods are applied with two objectives. First, these methods are useful for highlighting co-regulated gene clusters, i.e., clusters involved separately in the synthesis of different secondary metabolites but whose expressions are linked. For this purpose, the abovementioned bioinformatics tool for cluster detection, MIDASS, integrates an algorithm that highlights co-regulated clusters by comparing the transcriptomic data obtained under metabolite-producing and non-producing conditions<sup>107</sup>. More specifically, transcriptomics may be adapted to detect clusters containing genes located on different chromosomes. Since studies had demonstrated examples of this cross-chemistry between gene clusters<sup>184</sup>, Andersen *et al.* created a microarray compendium specific to *A. nidulans* from 44 culture experiments. Each clustering score was calculated to yield a numerical value with which to correlate the expression profiles of two genes. Using this method, the researchers highlighted a secondary metabolite,

nidulatin A, which involves genes from at least two chromosomes<sup>185</sup>. Nevertheless, RNA studies should include the detection of the corresponding secondary metabolites since many posttranscriptional regulations may disrupt the expression of a cluster, such as the posttranscriptional regulation of the synthesis of sterigmatocystin by PkaA in *A. nidulans*<sup>186</sup>.

Second, transcriptomic techniques could be applied to understand the links between genes and secondary metabolites or between genes and active properties. Thus, the clusters suggested by comparative phylogenomics to be responsible for the pathogenicity of *S. nodorum* were found to be specifically expressed during wheat infection based on comparison of *in planta* and *in vitro* microarrays<sup>187,188</sup>. Moreover, RNA-seq was implemented in studies of transcriptomic variations during the different stages of fungal infection<sup>188,189</sup>. An experiment involving a microarray comparison of *in planta* cultures led to the observation of a particular unknown cluster involved in wheat infection by *F. graminearum*<sup>190</sup>.

## 6. Modification of the Growth Environment: Environment-Dependent Approaches

To explain why 80% of the potential fungal secondary metabolites remain unknown, it was suggested that a majority of the gene clusters are silent under standard laboratory culture conditions<sup>79</sup>. Fungi modulate the synthesis of secondary metabolites depending on their needs for defense or colonization of the environment. This principle, commonly called OSMAC (one strain many compounds), indicates that every small change in the growth conditions can completely modify fungal secondary metabolism.



**Fig. 6** Illustration of the differences between natural growth conditions (sea or cereal crops for example) and standard laboratory growth conditions (liquid or solid media) for a fungus in terms of physical, chemical and interspecies cross-talk parameters. The slider positions reflect the complexity of the natural parameters involved in the production of the fungal secondary metabolome on the left side and their basic modulation under laboratory conditions on the right side. These changes in terms of quantity and type of the secondary

metabolite depending on the variation in each parameter are expressed under the general term OSMAC (One Strain-MANY Compounds).

In a review dedicated to bioactive fungal secondary metabolites, Bills *et al.* even argue that the lack of this knowledge was historically a major limiting factor in the discovery of fungal compounds of industrial interest<sup>11</sup>. Standard laboratory culture conditions are frequently arbitrarily monitored, and only several environmental factors, such as the temperature, nutrient availability or light exposure are controlled (Figure 6, right side). In contrast, natural growth conditions involve a complex balance between numerous variables that may fluctuate independently as well as the co-exposure to other fungi/bacteria (Figure 6, left side). Two different approaches to modify fungal growth conditions are considered: interspecies cross-talk and modification of the physicochemical culture conditions. Hereafter, we describe some studies in which modification of the growth environment was used as the strategy to discover fungal secondary metabolites. However, this method is commonly integrated as a complement to other strategies reported in this review.

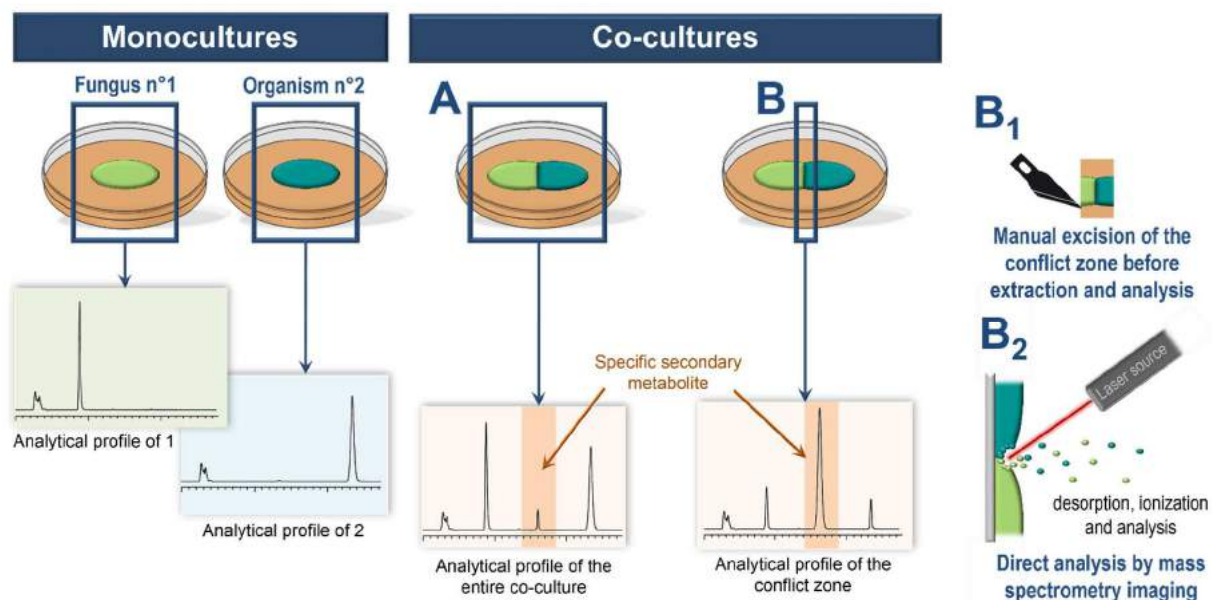
### 6.1 Interspecies Cross-Talk

The first approach for changing culture conditions involves the co-culturing of a fungus with other fungi or bacteria<sup>191</sup>. Interspecies cross-talk is used in research on antimicrobial compounds because co-culture seeks to stimulate the production of secondary metabolites involved in fungal defense against other microorganisms<sup>142,146,192</sup>. It was also reported that fungi might communicate with other microorganisms by integrating stimuli, for example, via epigenetic modifications<sup>193</sup>. As reported by Marmann *et al.*<sup>194</sup>, these strategies for natural product characterization can be based either on the analysis of the entire mixed fermentation or on the specific observation of the conflict zone.

In some studies, compounds were identified by comparing the analytical profiles of monoculture extracts with that of the entire co-culture extract<sup>195,196</sup>. Thus, a signal present in the co-culture but absent in the monocultures originates from secondary metabolites specifically synthesized in response to the conflict between the two organisms (Figure 7A). Zuck *et al.* characterized two xanthocillin analogues in a co-culture of *A. fumigatus* with a bacterial strain, one of which exhibited submicromolar activity against tumor cell lines<sup>197</sup>, and three lipopeptides were characterized<sup>198</sup>. Similarly, a co-culture of *F. tricinctum* and *Bacillus subtilis* led to the identification of two secondary metabolites: macrocarpon C and a derivative of anthranilic acid<sup>199</sup>. König *et al.* studied the influence of *Streptomyces rapamycinus* on the secondary metabolome of *A. fumigatus*. A group of unknown metabolites was highlighted, and fumicyclines A and B were characterized<sup>200</sup>. Through additional knock-out and epigenetic experiments, the authors have shown that the secretion of these molecules was stimulated by the bacteria through epigenetic modulation of the fungal genome. Since Chavez *et al.* emphasized the high potential of fungi living in harsh environments to produce

secondary metabolites with unusual chemical structures<sup>201</sup>, several studies have involved placing microorganisms from extreme environments into contact with one another. Using mixed fermentation, Zhu *et al.* characterized aspergicin from two marine-derived *Aspergillus* strains<sup>202</sup>, and Li *et al.* identified a xanthone derivative<sup>203</sup>. Similarly, the cytotoxic diterpenoids libertellenones A-D were characterized from *Libertella* sp.<sup>204</sup>, and the antibiotic pestalone was identified<sup>205</sup>. Oh *et al.* applied this method to 50 different marine fungi in co-culture with a

marine *Actinomycetes* species and characterized emerellamides A and B, with the former displaying low antimicrobial activity and weak cytotoxic activity against a tumor cell line<sup>206</sup>. Regarding other extreme environments, a co-culture of an *A. fumigatus* strain with a bacterium, both isolated from disused mines, led to the characterization of glionitrin A, an antitumor compound with submicromolar cytotoxic activity against four human cancer cell lines<sup>207</sup>.



**Fig. 7** Co-culturing of microorganisms for the identification of secondary metabolites. The signals specifically observed in co-culture (dark-orange image) originate from the secondary metabolites produced only via the communication between the two organisms. The analytical profile of the co-cultures may be obtained by (A) extracting the entire co-culture or (B) specifically analyzing the conflict zone. These profiles are generated by (B1) manual excision of the particular zone or (B2) direct analysis with MS imaging (here, MALDI-MSI).

## 6.2 Modulation of the Physicochemical Growth Conditions

The second environment-dependent strategy consists of abiotic modifications to the culture environment by modifying the light intensity and wavelength, carbon and nitrogen sources, ambient pH, redox status, culture configuration (static, shaken, liquid or solid) or many other factors<sup>142</sup>. Some studies have highlighted the modulation of fungal secondary metabolism merely according to the quality and source of water. Paranagama *et al.* characterized six fungal secondary metabolites simply by replacing tap water with mineral water in the culture medium<sup>208</sup>. The influence of salinity on secondary metabolomes was also explored. Marine-derived fungi are well known for their ability to produce numerous secondary metabolites<sup>209,210,211</sup>, and particularly antibacterial compounds<sup>212</sup>. An unknown compound with antibacterial activity was found to be produced by *S. elegans*, specifically in culture with 10% salinity<sup>213</sup>. The same experiment with *A. terreus* led to the characterization of three metabolites<sup>214</sup>.

Similarly, Zheng *et al.* observed the production of 11 unknown asprochracin-type cyclic tripeptides by growing *A. sclerotiorum* on hypersaline nutrient-limited media<sup>215</sup>.

The oxygen supply is another factor influencing the production of secondary metabolites. Comparing cultures of *Sphaeropsidales* sp. in shaken liquid media (i.e., sufficiently oxygenated cultures) and in static solid media (i.e., weakly oxygenated cultures), Bode *et al.* discovered eight spirobisnaphthalenes: cladospirones B to I<sup>216</sup>. In another study, the influence of aeration was analyzed, and 120 different cultures of *A. parasiticus* were assessed for antimicrobial activities. The association of bioguided purification and culture condition modulation revealed unknown secondary metabolites with bioactive properties<sup>31</sup>. In another study, the influence of temperature on the metabolome of *A. nidulans* was demonstrated. Among 45 tested conditions involving modulation of the culture temperature, oxygen supply and other parameters, growth at high temperature resulted in the production of two unknown natural compounds with weak antiproliferative effects, aspernidine A and B<sup>217</sup>.

The type and availability of the nutrient source were modulated in several studies. In some cases, the natural substrate allowed the secretion of specific secondary metabolites. For example, *Cephalosporium graminearum* was selected from soil for its antibiotic effect, but its active properties disappeared in a standard medium<sup>218</sup>. Similarly, Overy *et al.* characterized corymbiferone B, a compound only detected in *in planta* cultures of seven *Penicillium* ser. *Corymbifera* sp.<sup>219</sup>. After *in silico* analysis of the *A. nidulans* genome, Scherlack *et al.* noted some genes coding for anthranilate synthases known to synthesize alkaloid precursors. Since alkaloids were never detected from this fungus grown on standard laboratory medium, the fungus was grown on several cereals. Analysis of the extracts focused on aromatic nitrogen-containing metabolites, and four aspoquinolones A-D were characterized from the rice medium<sup>220</sup>. In contrast, the nutrient source can be deliberately limited. Sarkar *et al.* cultivated *A. nidulans* under continuous growth with limited supply rates of nitrate, orthophosphate and glucose. Under these conditions, the fungus synthesized a new molecule, preshamixanthone<sup>221</sup>. Similarly, sanghaspirodins A and B were discovered from *A. nidulans* grown with restricted nitrogen resources<sup>222</sup>. To address to the ecologically question related to determining the secondary metabolites produced by a fungal wheat pathogen, our laboratory used a protocol that combined a double isotope labeling and the OSMAC principle (see chapter 2 for details)<sup>45</sup>.

## 7. Conclusions

As described in this review, both targeted genome-based approaches and untargeted strategies have been used to characterize fungal secondary metabolites. The different strategies discussed in this review are schematized in Figure 1 and the resulting discoveries of fungal secondary metabolites are listed in Table 1. Genetic engineering may be applied for the discovery of secondary metabolites produced by orphan gene clusters, but several limitations remain, including the masking of unknown metabolites by the highly expressed metabolites and the silencing of the studied genes. These problems have been addressed through the development of techniques such as knock-down experiments, gene transfer to heterologous hosts and transcriptional activation. However, methods that generally require deletion, insertion or mutation of DNA sequences may be laborious in fungi because of their high frequency of non-homologous end joining which induces incorrect insertions of DNA sequences. A highly effective genome editing technology was recently established and has only begun to be adapted for fungal genetic engineering<sup>223</sup>. Discovered as an adaptive immunity of bacteria against viruses, the CRISPR-Cas9 system involves two components that induce double-strand breaks at specific loci of the DNA strand: a bacterial endonuclease Cas9 and a single-guide RNA (sgRNA) sequence that targets the DNA sequence. By simply designing specific sgRNA sequences, gene deletions, insertions or mutations can be precisely and efficiently induced in every organism<sup>224</sup>. In 2016, the first gene mutations, deletions and

replacements were performed in fungi using CRISPR-Cas9<sup>225,226,227,228</sup>, and this technology was reported in recent fungal papers as a very promising tool for genetic engineering<sup>229,230</sup> since it offers unparalleled potential for accurate and high throughput manipulations of fungal genomes.

As highlighted by the One Strain MAny Compounds principle, the observation of gene clusters enables the assessment of the entire potential of a fungus to generate secondary metabolites but does not represent what the fungus inherently expresses. Studies seeking to characterize naturally produced fungal compounds (i.e., in the context of toxicological investigations or studies of the communication between organisms) should consider the growth conditions and focus on the direct analysis of the secondary metabolites. The evolution of analytical instrumentation with higher resolution and sensitivity facilitates the development of untargeted approaches, particularly in the field of natural products research in which the molecules are produced in low quantities and as complex mixtures<sup>231,232</sup>. Microflow NMR is a promising tool that led to the discovery of several fungal products<sup>233,234</sup>. This high-sensitivity structural analysis of secondary metabolites is performed directly after chromatographic separation and avoids laborious purification steps. In a recent review, Covington *et al.* reported the various MS-based computational methods used to process metabolomic datasets and highlighted their application in the context of natural product research<sup>235</sup>. Aside from the technological advances, several methods, such as the SIL of fungal substrates, were elaborated to specifically target the secondary metabolites and simplify the data mining of complex datasets. New challenges associated with these untargeted metabolomics strategies lie in the development of computational tools for the mining and interpretation of data<sup>229</sup>.

Finally, the evolution of natural product research is moving toward an integrative approach that takes advantage of each tool: bioinformatics for elucidation of the fungal genomic potential, genome engineering for its targeted manipulation and metabolomics for analyzing the global composition of complex samples with high resolution and high sensitivity. These complementary processes were used together in a recent study of *Penicillium* sp. and led to the discovery of previously undescribed compounds<sup>236</sup>. Moreover, an extensive study combining genome mining, epigenetic techniques, transcriptomics, proteomics, and metabolomics that also took into account the fungal culture conditions characterized two novel secondary metabolites as virulent factors influencing the destruction of rice fields by *F. fujikoro*<sup>237</sup>. However, the development of multiomics strategies produces big data and the problem of data integration. The continued improvement of the data mining to handle the huge quantities of data, produced in this research and the ability to share data on unidentified compounds with other researchers are crucial pursuits. Such integrative studies also necessitate the development of bioinformatics tools to process together all different omics data. Finally, efforts must also be focused on the accessibility and comparability of fungal omics datasets as well as the standardization of the analytical quality<sup>229</sup>.

Strategy (bold) and additional method(s)	Metabolite(s) discovered	Fungal species	References
<b>TARGETED PURIFICATION</b>			
<b>Purification of crude extract</b>			
	Arugosins H and G	<i>Emericella nidulans</i>	12
	Trichodermamides A and B	<i>Trichoderma virens</i>	13
	Epicoccamides B-D	<i>Epicoccum</i> sp.	14
	Isoharzianic acid	<i>Trichoderma harzianum</i>	15
	Variecolorins A-L	<i>Aspergillus variecolor</i>	16
+ Computational analysis with MeHaloCoA	Griseophenone I Chlorogriseofulvin	<i>Penicillium canescens</i>	17
<b>Bioguided purification</b>			
	Fumonisin B1 and B2	<i>Fusarium moniliforme</i>	19
	Mevinolin	<i>Aspergillus terreus</i>	20
	Fusamine Fusarine And 3 others	<i>Fusarium incarnatum</i>	21
	Besudon	<i>Mollisia benesuada</i>	22
	Epolones A and B, Pycnidione Coniothyrione	<i>Neosetophoma samarorum</i>	23
	Trichodermamide A Espergillazine A	<i>Spicaria elegans</i>	24
	Cytochalasins Z <sub>10</sub> -Z <sub>15</sub>	<i>Spicaria elegans</i>	25
	Variecolortides A to C	<i>Aspergillus variecolor</i>	26
	7 new communesins	<i>Penicillium expansum</i>	27
	Sorokiniol	<i>Bipolaris sorokiniana</i>	28
	Ligerin	<i>Penicillium</i> strain from sect. <i>Canescentia</i>	29
	Pannomycin	<i>Geomyces pannorum</i>	30
<b>Untargeted metabolomics</b>			
+ Stable Isotope Labeling	9 new compounds	<i>Fusarium graminearum</i>	44
+ Stable Isotope Labeling	A new fumigaclavine analogue	<i>Aspergillus fumigatus</i>	45
+ Stable Isotope Labeling + Computational analysis with molecular networks	4 analogues of nidulanins and fungisporins	<i>Aspergillus nidulans</i>	37
+ Stable Isotope Labeling + Adaptation of the growth environment + Computational analysis with molecular networks	82 new secondary metabolites	<i>Penicillium verrucosum</i>	46
<b>GENOME MINING</b>			
<b>Gene knock-out</b>			
	Emericellamides C-F	<i>Aspergillus nidulans</i>	127
	Aspercryptin	<i>Aspergillus nidulans</i>	132
<b>Heterologous host transfer</b>			
	Neosartoricin B	<i>Trichophyton tonsurans</i>	139
	Terezine D	<i>Aspergillus terreus</i>	144
+ Gene knock-out + Specific pathway overexpression	Asperfuranone	<i>Aspergillus terreus</i>	136
+ Stable isotope labeling	Aculins A and B Epi-aculin A	<i>Aspergillus aculeatus</i>	140
<b>Specific pathway overexpression</b>			
	Aspyridones A and B	<i>Aspergillus nidulans</i>	149

	Fusarielins, F, G and H	<i>Fusarium graminearum</i>	150
	Neosartoricin	<i>Aspergillus fumigatus</i> and <i>Neosartorya fischeri</i>	151
	Apicidin F	<i>Fusarium fujikuroi</i>	152
	7 new metabolites derived from the $\beta$ -resorcyclaldehyde	<i>Aspergillus nidulans</i>	153
	Microperfuraneone	<i>Aspergillus nidulans</i>	154
+ Gene knock-out	Asperfuraneone	<i>Aspergillus nidulans</i>	133
<b>Global transcription modulation</b>			
	Terrequinone A	<i>Aspergillus nidulans</i>	157
<b>Mutagenesis strategies</b>			
	Apicidin E	<i>Fusarium semitectum</i>	171
	Purpurogemutantin Purpurogemutantidin	<i>Penicillium purpurogenum</i>	32
	Penicimutanolone Penicimutanin A and B Penicimutatin	<i>Penicillium purpurogenum</i>	164
	Antineocyclocitrinols A and B 23-O-methylantineocyclocitrinol	<i>Penicillium purpurogenum</i>	165
	Epiremisorpine B and B1 Isoconiochaetone C	<i>Penicillium purpurogenum</i>	166
	Janthinone Fructigenine A Aspterric acid methyl ester	<i>Penicillium purpurogenum</i>	167
	Curvularin Penicitrinone A Erythro-23-O-methylneocyclocitrinol 22E-7 $\alpha$ -methoxy-5 $\alpha$ ,6 $\alpha$ -epoxyergosta-8(14),22-dien-3 $\beta$ -ol	<i>Penicillium purpurogenum</i>	168
	Penicimutide Cyclo(L-Val-L-Pro) Cyclo(L-Ile-L-Pro) Cyclo(L-Leu-L-Pro) Cyclo(L-Phe-L-Pro)	<i>Penicillium purpurogenum</i>	169
	Phenethyl 5-oxo-L-prolinate Cyclo(D-Pro-D-Phe) Cyclo(D-Tyr-D-Pro) 3 $\beta$ ,5 $\alpha$ ,9 $\alpha$ -trihydroxy-(22E,24R)-ergosta-7,22-dien-6-one	<i>Aspergillus versicolor</i>	170
	Mutolide	<i>Sphaeropsidales</i> sp.	162
<b>Epigenetic</b>			
	(7S)-(+)-7-O-methylsydonol (7S,11S)-(+)-12-hydroxysydonic acid 7-deoxy-7,14-didehydroxydonol (S)-(+)-sydonol	<i>Aspergillus sydowii</i>	177
	Nigeron A	<i>Aspergillus niger</i>	182
	3 oxylipins Cladochromes F and G	<i>Cladosporium cladosporioides</i>	178
	Lunalides A and B	<i>Diatrype</i> sp.	178
	Atlantinones A and B	<i>Penicillium citreonigrum</i>	179
+ Heterologous host transfer + Transcriptomics	(1R,4R,5S)-guaia-6,10(14)-diene	<i>Fusarium fujikuroi</i>	181
+ Untargeted metabolomics	Fellutamide B Antibiotic 1656-G Antibiotic 3127	<i>Aspergillus nidulans</i>	183

<b>TRANSCRIPTOMICS</b>			
	Nidulanin A	<i>Aspergillus nidulans</i>	185
<b>MODIFICATION OF GROWTH ENVIRONMENT</b>			
<b>Interspecies cross-talk</b>			
+ Bioguided purification	Pestalone	<i>Pestalotia</i> (CNL-365) (+ bacterial strain CNJ-328)	205
+ Bioguided purification	4 Mycoalexins	<i>Eutypa lata</i> (+ <i>Botryosphaeria obtuse</i> )	234
+ Purification of crude extract	Secopenicillide C	<i>Penicillium pinophilum</i> (+ <i>Trichoderma harzianum</i> )	195
+ Purification of crude extract	Subenniatis A and B	<i>Fusarium tricinctum</i> (+ <i>Fusarium begonia</i> )	196
+ Purification of crude extract	Macrocapon C And 1 derived of anthranilic acid	<i>Fusarium tricinctum</i> (+ <i>Bacillus subtilis</i> )	199
+ Purification of crude extract	4''-hydroxysulfoxy-2,2''-dimethylthielavin P Its nonsulfated form And 3 other related compounds	<i>Trichophyton rubrum</i> (+ <i>Bionectria ochroleuca</i> )	238
+ Purification of crude extract	Fumiformamide N,N <sub>0</sub> -((1Z,3Z)-1,4-bis(4-methoxyphenyl)buta- 1,3-diene-2,3-diyl)diformamide	<i>Aspergillus fumigatus</i> (+ <i>Streptomyces peucetius</i> )	197
+ Purification of crude extract	Aspergicin	<i>Aspergillus</i> sp. (+ another <i>Aspergillus</i> sp.)	202
+ Purification of crude extract	8-hydroxy-3-methyl-9-oxo-9H-xanthene-1- carboxylic acid methyl ether	Strain No. K38 (+ strain No. E33)	203
+ Purification of crude extract	Libertellenone A-D	<i>Libertella</i> sp. (+ a marine $\alpha$ -proteobacterium)	204
	Acremostatins A, B and C	<i>Acremonium</i> sp. (+ <i>Mycogone rosea</i> )	198
+ Purification of crude extract	Emericellamides A and B	<i>Emericella</i> sp. (+ <i>Salinispora arenicola</i> )	206
+ Purification of crude extract	Glionitrin A	<i>Aspergillus fumigatus</i> (+ a <i>Sphingomonas</i> bacterial strain)	207
+ Purification of crude extract	Fumicycline A and B	<i>Aspergillus fumigatus</i> (+ <i>Streptomyces rapamycinus</i> )	200
+ Purification of crude extract + Genome mining + Transcriptomics			
+ Computational analysis with molecular networks	2 dimeric phenazines And other bioconversions of bacterial compounds by the fungus	<i>Aspergillus fumigatus</i> (+ <i>Pseudomonas aeruginosa</i> )	239
<b>Modification of physicochemical growth conditions</b>			
+ Bioguided purification	New antibiotics	<i>Aspergillus parasiticus</i>	31
+ Purification of crude extract	Cytosporones F-I Quadrisepin A 5'-hydroxymonocillin III	<i>Paraphaeosphaeria quadrisepinata</i>	208
+ Purification of crude extract	(2E,2'Z)-3,3'-(6,6'-dihydroxybiphenyl-3,3'- diyl)diacrylic acid	<i>Spicaria elegans</i>	213
+ Purification of crude extract	Terremides A and B Terrelactone A	<i>Aspergillus terreus</i>	214
+ Purification of crude extract	Sclerotiotides A-K	<i>Aspergillus sclerotiorum</i>	215
+ Purification of crude extract	Cladospirones B-I	<i>Sphaeropsidales</i> sp.	216
+ Purification of crude extract	Aspernidines A and B	<i>Aspergillus nidulans</i>	217
+ Purification of crude extract	Corymbiferone B	7 <i>Penicillium</i> ser. <i>Corymbifera</i>	219
+ Purification of crude extract	Aspoquinolones A-D	<i>Aspergillus nidulans</i>	220
+ Purification of crude extract + Transcriptomics	Pre-shamixanthone	<i>Aspergillus nidulans</i>	221
+ Purification of crude extract	Sanghaspirodins A and B	<i>Aspergillus nidulans</i>	222



+ Transcriptomics + Genome mining			
+ Transcriptomics + Genome mining + Proteomics + Phylogenetics	4 new metabolites	<i>Fusarium fujikoro</i>	237

**Table 1.** Strategies used for fungal secondary metabolite discovery.

## 8. Acknowledgements

The platform MetaToul-AXIOM is integrated in the French national infrastructure of metabolomics and fluxomics: MetaboHUB (MetaboHUB-ANR-11-INBS-0010). This study was co-funded by INRA and French Minister of Higher Education and Research in the context of a project supported by French National Agency of Research (ANR-15-CE21-0010-21 Newmyco). TH was supported by a Fellowship from the Ministère de l'Éducation Nationale, de la Recherche et de la Technologie. The authors are grateful to Dr. Rougeulle from the University Paris Diderot (CEDC/CNRS, Paris, France) for her kind help and her expertise in genetics and epigenetics.

## 9. References

- D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2006, **70**, 461–477.
- A. L. Demain and A. Fang, *The natural functions of secondary metabolites*, Springer B., 2000.
- R. Hare, W. Florey, A. D. Gardner, N. G. Heatley, M. A. Jennings, J. Orr-Ewing, A. G. Sanders, E. P. Abraham, E. Chain, C. M. Fletcher and H. W. Florey, *Med. Hist.*, 1982, **26**, 1–24.
- D. W. Denning and M. J. Bromley, *Science (80-. )*, 2015, **347**, 1414–1416.
- G. C. Merhoff and J. M. Porter, *Ann. Surg.*, 1974, **180**, 773.
- N. Bradburn, R. D. Coker and G. Blunden, *Phytochemistry*, 1994, **35**, 817.
- S. Brase, A. Encinas, J. Keck and C. F. Nising, *Chem. Rev.*, 2009, **109**, 3903–3990.
- J. W. Bennett and M. Klich, *Clin. Microbiol. Rev.*, 2003, **16**, 497–516.
- M. C. Fisher, D. A. Henk, C. J. Briggs, J. S. Brownstein, L. C. Madoff, S. L. McCraw and S. J. Gurr, *Nature*, 2013, **484**, 1–18.
- M. Blackwell, *Am. J. Bot.*, 2011, **98**, 426–438.
- G. F. Bills and J. B. Gloer, *Microbiol Spectr.*, 2016, **4**.
- A. Kralj, S. Kehraus, A. Krick, E. Eguereva, G. Kelter, M. Maurer, A. Wortmann, H. H. Fiebig and G. M. König, *J. Nat. Prod.*, 2006, **69**, 995–1000.
- E. Garo, C. M. Starks, P. R. Jensen, W. Fenical, E. Lobkovsky and J. Clardy, *J. Nat. Prod.*, 2003, **66**, 423–426.
- H. V. K. Wangun, H.-M. Dahse and C. Hertweck, *J. Nat. Prod.*, 2007, **70**, 1800–1803.
- F. Vinale, G. Manganiello, M. Nigro, P. Mazzei, A. Piccolo, A. Pascale, M. Ruocco, R. Marra, N. Lombardi, S. Lanzuise, R. Varlese, P. Cavallo, M. Lorito and S. L. Woo, *Molecules*, 2014, **19**, 9760–9772.
- W.-L. Wang, Z.-Y. Lu, H.-W. Tao, T.-J. Zhu, Y.-C. Fang, Q.-Q. Gu and W.-M. Zhu, *J. Nat. Prod.*, 2007, **70**, 1558–1564.
- C. Roullier, Y. Guitton, M. Valery, S. Amand, S. Prado, T. Robiou du Pont, O. Grovel and Y. F. Pouchus, *Anal. Chem.*, 2016, **88**, 9143–9150.
- G. W. Gribble, *Mar. Drugs*, 2015, **13**, 4044–4136.
- W. C. A. Gelderblom, K. Jaskiewicz, W. F. O. Marasas, P. G. Thiel, R. M. Horak, R. Vleggaar and N. P. J. Kriek, *Appl. Environ. Microbiol.*, 1988, **54**, 1806–1811.
- A. W. Alberts, J. Chen, G. Kuron, V. Hunt, J. Huff, C. Hoffman, J. Rothrock, M. Lopez, H. Joshua, E. Harris, A. Patchett, R. Monaghan, S. Currie, E. Stapley, G. Albers-Schonberg, O. Hensens, J. Hirshfield, K. Hoogsteen, J. Liesch and J. Springer, *PNAS*, 1980, **77**, 3957–3961.
- L. Ding, H. M. Dahse and C. Hertweck, *J. Nat. Prod.*, 2012, **75**, 617–621.
- E. Thines, W.-R. Arendholz, H. Anke and O. Sterner, *J. Antibiot. (Tokyo)*, 1997, **50**, 13–17.
- D. P. Overy, F. Berrue, H. Correa, N. Hanif, K. Hay, M. Lanteigne, K. Mquilian, S. Duffy, P. Boland, R. Jagannathan, G. S. Carr, M. Vansteeland and R. G. Kerr, *Mycology*, 2014, **5**, 130–144.
- R. Liu, Q.-Q. Gu, W.-M. Zhu, C.-B. Cui and G.-T. Fan, *Arch Pharm Res*, 2005, **28**, 1042–1046.
- R. Liu, Z. Lin, T. Zhu, Y. Fang, Q. Gu and W. Zhu, *J. Nat. Prod.*, 2008, **71**, 1127–1132.
- W. L. Wang, T. J. Zhu, H. W. Tao, Z. Y. Lu, Y. C. Fang, Q. Q. Gu and W. M. Zhu, *Chem. Biodivers.*, 2007, **4**, 2913–2919.
- I. Kerzaon, Y. F. Pouchus, F. Monteau, B. Le Bizec, M. R. Nourrisson, J. F. Biard and O. Grovel, *Rapid Commun. Mass Spectrom.*, 2009, **23**, 3928–3938.
- L. Ali, A. L. Khan, J. Hussain, A. Al-Harrasi, M. Waqas, S.-M. Kang, A. Al-Rawahi and I.-J. Lee, *BMC Microbiol.*, 2016, **16**, 1.
- M. Vansteelandt, E. Blanchet, M. Egorov, F. Petit, L. Toupet, A. Bondon, F. Monteau, B. Bizec, O. P. Thomas, Y. F. Pouchus, R. Le Bot and O. Grovel, *J. Nat. Prod.*, 2013, **76**, 297–301.
- C. A. Parish, M. De La Cruz, S. K. Smith, D. Zink, J. Baxter, S.

- Tucker-Samaras, J. Collado, G. Platas, G. Bills, M. T. Diez, F. Vicente, F. Pelaez and K. Wilson, *J. Nat. Prod.*, 2009, **72**, 59–62.
- 31 A. A. P. Bracarense and J. A. Takahashi, *Brazilian J. Microbiol.*, 2014, **45**, 313–321.
- 32 S. M. Fang, C. Bin Cui, C. W. Li, C. J. Wu, Z. J. Zhang, L. Li, X. J. Huang and W. C. Ye, *Mar. Drugs*, 2012, **10**, 1266–1287.
- 33 J. Smedsgaard and J. Nielsen, *J. Exp. Bot.*, 2005, **56**, 273–286.
- 34 S. Neumann and S. Böcker, *Anal. Bioanal. Chem.*, 2010, **398**, 2779–2788.
- 35 T. Kind and O. Fiehn, *Bioanal. Rev.*, 2010, **2**, 23–60.
- 36 J.-L. Wolfender, G. Marti, A. Thomas and S. Bertrand, *J. Chromatogr. A*, 2015, **1382**, 136–164.
- 37 A. Klitgaard, J. B. Nielsen, R. J. N. Frandsen, M. R. Andersen and K. F. Nielsen, *Anal. Chem.*, 2015, **87**, 6520–6526.
- 38 C. Bueschl, B. Kluger, M. Lemmens, G. Adam, G. Wiesenberger, V. Maschietto, A. Marocco, J. Strauss, S. Bödi, G. G. Thallinger, R. Krska and R. Schuhmacher, *Metabolomics*, 2014, **10**, 754–769.
- 39 C. Bueschl, R. Krska, B. Kluger and R. Schuhmacher, *Anal. Bioanal. Chem.*, 2013, **405**, 27–33.
- 40 F. A. de Jong and C. Beecher, *Bioanalysis*, 2012, **4**, 2303–2314.
- 41 C. S. Clendinen, G. S. Stupp, R. Ajredini, B. Lee-McMullen, C. Beecher and A. S. Edison, *Front. Plant Sci.*, 2015, **6**, 611.
- 42 C. Bueschl, B. Kluger, F. Berthiller, G. Lirk, S. Winkler, R. Krska and R. Schuhmacher, *Bioinformatics*, 2012, **28**, 736–738.
- 43 X. Huang, Y.-J. Chen, K. Cho, I. Nikolskiy, P. A. Crawford and G. J. Patti, *Anal. Chem.*, 2014, **86**, 1632–1639.
- 44 N. K. N. Neumann, S. M. Lehner, B. Kluger, C. Bueschl, K. Sedelmaier, M. Lemmens, R. Krska and R. Schuhmacher, *Anal. Chem.*, 2014, **86**, 7320–7327.
- 45 P. M. Cano, E. L. Jamin, S. Tadrist, P. Bourdaud’Hui, M. Péan, L. Debrauwer, I. P. Oswald, M. Delaforge and O. Puel, *Anal. Chem.*, 2013, **85**, 8412–8420.
- 46 T. Hautbergue, O. Puel, S. Tadrist, L. Meneguetti, M. Péan, M. Delaforge, L. Debrauwer, I. P. Oswald and E. L. Jamin, *J. Chrom. B. Chrom. B.*, 2017, **DOI:10.101**.
- 47 M. T. Henke and N. L. Kelleher, *Nat. Prod. Rep.*, 2016, **33**, 942–950.
- 48 K. F. Nielsen, T. O. Larsen, N. Keller and T. S. Bugni, *Front. Microbiol.*, 2015, **6**, 71.
- 49 S. P. Gaudêncio and F. Pereira, *Nat. Prod. Rep.*, 2015, **32**, 779–810.
- 50 H. Laatsch, 2012.
- 51 Chapman, 2005.
- 52 T. El-Elimat, M. Figueroa, B. M. Ehrmann, N. B. Cech, C. J. Pearce and N. H. Oberlies, *J. Nat. Prod.*, 2013, **76**, 1709–1716.
- 53 K. F. Nielsen and J. Smedsgaard, *J. Chromatogr. A*, 2003, **1002**, 111–136.
- 54 S. Kildgaard, M. Mansson, I. Dosen, A. Klitgaard, J. C. Frisvad, T. O. Larsen and K. F. Nielsen, *Mar. Drugs*, 2014, **12**, 3681–3705.
- 55 H. Horai, M. Arita, S. Kanaya, Y. Nihei, T. Ikeda, K. Suwa, Y. Ojima, K. Tanaka, S. Tanaka, K. Aoshima, Y. Oda, Y. Kakazu, M. Kusano, T. Tohge, F. Matsuda, Y. Sawada, M. Y. Hirai, H. Nakanishi, K. Ikeda, N. Akimoto, T. Maoka, H. Takahashi, T. Ara, N. Sakurai, H. Suzuki, D. Shibata, S. Neumann, T. Iida, K. Tanaka, K. Funatsu, F. Matsuura, T. Soga, R. Taguchi, K. Saito and T. Nishioka, *J. Mass Spectrom.*, 2010, **45**, 703–714.
- 56 Y. Sawada, R. Nakabayashi, Y. Yamada, M. Suzuki, M. Sato, A. Sakata, K. Akiyama, T. Sakurai, F. Matsuda, T. Aoki, M. Y. Hirai and K. Saito, *Phytochemistry*, 2012, **82**, 38–45.
- 57 J. Y. Yang, L. M. Sanchez, C. M. Rath, X. Liu, P. D. Boudreau, N. Bruns, E. Glukhov, A. Wodtke, R. De Felicio, A. Fenner, W. R. Wong, R. G. Linington, L. Zhang, H. M. Debonsi, W. H. Gerwick and P. C. Dorrestein, *J. Nat. Prod.*, 2013, **76**, 1686–1699.
- 58 M. Wang, J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. Kaponov, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A. V. Melnik, M. J. Meehan, W.-T. Liu, M. Crüsemann, P. D. Boudreau, E. Esquenazi, M. Sandoval-Calderón, R. D. Kersten, L. A. Pace, R. A. Quinn, K. R. Duncan, C.-C. Hsu, D. J. Floros, R. G. Gavilan, K. Kleigrew, T. Northen, R. J. Dutton, D. Parrot, E. E. Carlson, B. Aigle, C. F. Michelsen, L. Jelsbak, C. Sohlenkamp, P. Pevzner, A. Edlund, J. McLean, J. Piel, B. T. Murphy, L. Gerwick, C.-C. Liaw, Y.-L. Yang, H.-U. Humpfer, M. Maansson, R. A. Keyzers, A. C. Sims, A. R. Johnson, A. M. Sidebottom, B. E. Sedio, A. Klitgaard, C. B. Larson, C. A. Boya P, D. Torres-Mendoza, D. J. Gonzalez, D. B. Silva, L. M. Marques, D. P. Demarque, E. Pociute, E. C. O’Neill, E. Briand, E. J. N. Helfrich, E. A. Granatosky, E. Glukhov, F. Ryffel, H. Houson, H. Mohimani, J. J. Kharbush, Y. Zeng, J. A. Vorholt, K. L. Kurita, P. Charusanti, K. L. McPhail, K. F. Nielsen, L. Vuong, M. Elfeki, M. F. Traxler, N. Engene, N. Koyama, O. B. Vining, R. Baric, R. R. Silva, S. J. Mascuch, S. Tomasi, S. Jenkins, V. Macherla, T. Hoffman, V. Agarwal, P. G. Williams, J. Dai, R. Neupane, J. Gurr, A. M. C. Rodríguez, A. Lamsa, C. Zhang, K. Dorrestein, B. M. Duggan, J. Almaliti, P.-M. Allard, P. Phapale, L.-F. Nothias, T. Alexandrov, M. Litaudon, J.-L. Wolfender, J. E. Kyle, T. O. Metz, T. Peryea, D.-T. Nguyen, D. VanLeer, P. Shinn, A. Jadhav, R. Müller, K. M. Waters, W. Shi, X. Liu, L. Zhang, R. Knight, P. R. Jensen, B. Ø. Palsson, K. Pogliano, R. G. Linington, M. Gutiérrez, N. P. Lopes, W. H. Gerwick, B. S. Moore, P. C. Dorrestein and N. Bandeira, *Nat. Biotechnol.*, 2016, **34**, 828–837.
- 60 H. Mohimani, A. Gurevich, A. Mikheenko, N. Garg, L.-F. Nothias, A. Ninomiya, K. Takada, P. C. Dorrestein and P. A. Pevzner, *Nat. Chem. Biol.*, 2016, **13**, 30–37.
- 61 J. Ng, N. Bandeira, W. Liu, M. Ghassemian, T. L. Simmons, W. Gerwick, R. Linington, P. Dorrestein and P. Pevzner, *Nat. Methods*, 2009, **6**, 596–599.
- 62 A. Ibrahim, L. Yang, C. Johnston, X. Liu, B. Ma and N. a Magarvey, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 19196–

- 201.
- 63 W. Demuth, M. Karlovits and K. Varmuza, *Anal. Chim. Acta*, 2004, **516**, 75–85.
- 64 H. Mohimani and P. A. Pevzner, *Nat. Prod. Rep.*, 2016, **33**, 73–86.
- 65 J. J. J. van der Hooft, J. Wandy, M. P. Barrett, K. E. V Burgess and S. Rogers, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 13738–13743.
- 66 S. Böcker and K. Dührkop, *J. Cheminform.*, 2016, **8**, 1–26.
- 67 A. Vaniya and O. Fiehn, *TrAC Trends Anal. Chem.*, 2015, **69**, 52–61.
- 68 F. Rasche, A. Svatos, R. K. Maddula, C. Boettcher and S. Boecker, *Anal. Chem.*, 2011, **83**, 1243–1251.
- 69 F. Rasche, K. Scheubert, F. Hufsky, T. Zichner, M. Kai, A. Svatos and S. Bocker, *Anal. Chem.*, 2012, **84**, 3417–3426.
- 70 H. Mohimani, W.-T. Liu, R. D. Kersten, B. S. Moore, P. C. Dorrestein and P. A. Pevzner, *J. Nat. Prod.*, 2014, **77**, 1902–1909.
- 71 T. H. J. Niedermeyer and M. Strohal, *PLoS One*, 2012, **7**, e44913.
- 72 K. Suhre and P. Schmitt-Kopplin, *Nucleic Acids Res.*, 2008, **36**, 481–484.
- 73 K. Dührkop, H. Shen, M. Meusel, J. Rousu and S. Böcker, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 12580–5.
- 74 F. Allen, A. Pon, M. Wilson, R. Greiner and D. Wishart, *Nucleic Acids Res.*, 2014, **42**, 1–6.
- 75 F. Allen, R. Greiner and D. Wishart, *Metabolomics*, 2015, **11**, 98–110.
- 76 L. Ridder, J. van der Hooft, S. Verhoeven, R. C. H. de Vos, R. J. Bino and J. Vervoort, *Anal. Chem.*, 2013, **85**, 6033–6040.
- 77 S. Wolf, S. Schmidt, M. Müller-Hannemann and S. Neumann, *BMC Bioinformatics*, 2010, **11**, 148.
- 78 C. L. Schardl, *Chem. Biol.*, 2006, **13**, 5–6.
- 79 P. Wiemann and N. P. Keller, *J. Ind. Microbiol. Biotechnol.*, 2014, **41**, 301–313.
- 80 A. A. Brakhage and V. Schroeckh, *Fungal Genet. Biol.*, 2011, **48**, 15–22.
- 81 R. A. Owens, S. Hammel, K. J. Sheridan, G. W. Jones and S. Doyle, *PLoS One*, 2014, **9**, e106942.
- 82 S. B. Bumpus, B. S. Evans, P. M. Thomas, I. Ntai and N. L. Kelleher, *Nat. Biotechnol.*, 2009, **27**, 951–956.
- 83 J. L. Meier and M. D. Burkart, *Curr. Opin. Chem. Biol.*, 2011, **15**, 48–56.
- 84 J. L. López, *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.*, 2007, **849**, 190–202.
- 85 S. Wiese, K. A. Reidegeld, H. E. Meyer and B. Warscheid, *Proteomics*, 2007, **7**, 340–350.
- 86 R. D. Taylor, A. Sapano, B. Blackwell, V. Anoop, S. Gleddie, N. A. Tinker and L. J. Harris, *Proteomics*, 2008, **8**, 2256–2265.
- 87 J. L. Meier, S. Niessen, H. S. Hoover, T. L. Foley, B. F. Cravatt and M. D. Burkart, *ACS Chem. Biol.*, 2009, **4**, 948–957.
- 88 J. L. Meier, A. C. Mercer and M. D. Burkart, *J. Am. Chem. Soc.*, 2008, **130**, 5443–5445.
- 89 M. Machida, K. Asai, M. Sano, T. Tanaka, T. Kumagai, G. Terai, K. Kusumoto, T. Arima, O. Akita, Y. Kashiwagi, K. Abe, K. Gomi, H. Horiuchi, K. Kitamoto, T. Kobayashi, M. Takeuchi, D. W. Denning, J. E. Galagan, W. C. Nierman, J. Yu, D. B. Archer, J. W. Bennett, D. Bhatnagar, T. E. Cleveland, N. D. Fedorova, O. Gotoh, H. Horikawa, A. Hosoyama, M. Ichinomiya, R. Igarashi, K. Iwashita, P. R. Juvvadi, M. Kato, Y. Kato, T. Kin, A. Kokubun, H. Maeda, N. Maeyama, J. Maruyama, H. Nagasaki, T. Nakajima, K. Oda, K. Okada, I. Paulsen, K. Sakamoto, T. Sawano, M. Takahashi, K. Takase, Y. Terabayashi, J. R. Wortman, O. Yamada, Y. Yamagata, H. Anazawa, Y. Hata, Y. Koide, T. Komori, Y. Koyama, T. Minetoki, S. Suharnan, A. Tanaka, K. Isono, S. Kuhara, N. Ogasawara and H. Kikuchi, *Nature*, 2005, **438**, 1157–1161.
- 90 H. Gross, *Appl. Microbiol. Biotechnol.*, 2007, **75**, 267–277.
- 91 M. Zerikly and G. L. Challis, *ChemBioChem*, 2009, **10**, 625–633.
- 92 J. M. Winter, S. Behnken and C. Hertweck, *Curr. Opin. Chem. Biol.*, 2011, **15**, 22–31.
- 93 K. Ochi, *J. Antibiot. (Tokyo)*, 2016, 1–16.
- 94 N. Ziemert, M. Alanjary and T. Weber, *Nat. Prod. Rep.*, 2016, **33**, 988–1005.
- 95 K. Ochi and T. Hosaka, *Appl. Microbiol. Biotechnol.*, 2013, **97**, 87–98.
- 96 T. Weber and H. U. Kim, *Synth. Syst. Biotechnol.*, 2016, **1**, 69–79.
- 97 T. A. J. Van der Lee and M. H. Medema, *Fungal Genet. Biol.*, 2016, **89**, 29–36.
- 98 A. K. Chavali and S. Y. Rhee, *Brief. Bioinform.*, 2017, 1–13.
- 99 N. Khaldi, F. T. Seifuddin, G. Turner, D. Haft, W. C. Nierman, K. H. Wolfe and N. D. Fedorova, *Fungal Genet. Biol.*, 2010, **47**, 736–741.
- 100 A. Starcevic, J. Zucko, J. Simunkovic, P. F. Long, J. Cullum and D. Hranueli, *Nucleic Acids Res.*, 2008, **36**, 6882–6892.
- 101 M. H. Medema, K. Blin, P. Cimermancic, V. De Jager, P. Zakrzewski, M. A. Fischbach, T. Weber, E. Takano and R. Breitling, *Nucleic Acids Res.*, 2011, **39**, W339–346.
- 102 K. Blin, M. H. Medema, D. Kazempour, M. A. Fischbach, R. Breitling, E. Takano and T. Weber, *Nucleic Acids Res.*, 2013, **Jul**, gkt449.
- 103 T. Weber, K. Blin, S. Duddela, D. Krug, H. U. Kim, R. Brucoleri, S. Y. Lee, M. A. Fischbach, R. Müller, W. Wohlleben, R. Breitling, E. Takano and M. H. Medema, *Nucleic Acids Res.*, 2015, **43**, W237–43.
- 104 M. A. Skinnider, C. A. DeJong, P. N. Rees, C. W. Johnston, H. Li, A. L. H. Webster, M. A. Wyatt and N. A. Magarvey, *Nucleic Acids Res.*, 2015, **43**, 9645–9662.
- 105 T. Wolf, V. Shelest, N. Nath and E. Shelest, *Bioinformatics*, 2016, **32**, 1138–1143.
- 106 I. Takeda, M. Umemura, H. Koike, K. Asai and M. Machida, *DNA Res.*, 2014, **21**, 447–457.
- 107 M. Umemura, H. Koike, N. Nagano, T. Ishii, J. Kawano, N. Yamane, I. Kozono, K. Horimoto, K. Shin-ya, K. Asai, J. Yu, J. W. Bennett and M. Machida, *PLoS One*, 2013, **8**, e84028.

- 108 M. Umemura, H. Koike and M. Machida, *Front. Microbiol.*, 2015, **6**, 371.
- 109 P. Wiemann, C.-J. Guo, J. M. Palmer, R. Sekonyela, C. C. C. Wang and N. P. Keller, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 17065–70.
- 110 T. J. Sharpton, J. E. Stajich, S. D. Rounsley, M. J. Gardner, J. R. Wortman, V. S. Jordan, R. Maiti, C. D. Kodira, D. E. Neafsey, Q. Zeng, C.-H. Hung, C. McMahan, A. Muszewska, M. Grynberg, M. A. Mandel, E. M. Kellner, B. M. Barker, J. N. Galagiani, M. J. Orbach, T. N. Kirkland, G. T. Cole, M. R. Henn, B. W. Birren and J. W. Taylor, *Genome Res.*, 2009, **19**, 1722.
- 111 R. A. Syme, J. K. Hane, T. L. Friesen and R. P. Oliver, *G3 Genes/ Genomes/ Genet.*, 2013, **3**, 959–969.
- 112 B. J. Condon, Y. Leng, D. Wu, K. E. Bushley, R. A. Ohm, R. Otilar, J. Martin, W. Schackwitz, J. Grimwood, N. A. I. MohdZainudin, C. Xue, R. Wang, V. A. Manning, B. Dhillon, Z. J. Tu, B. J. Steffenson, A. Salamov, H. Sun, S. Lowry, K. LaButti, J. Han, A. Copeland, E. Lindquist, K. Barry, J. Schmutz, S. E. Baker, L. M. Ciuffetti, I. V. Grigoriev, S. Zhong and B. G. Turgeon, *PLoS Genet.*, 2013, **9**, e1003233.
- 113 K. E. Bushley and B. G. Turgeon, *BMC Evol. Biol.*, 2010, **10**, 26.
- 114 Y. H. Chooi and Y. Tang, *J. Org. Chem.*, 2012, **77**, 9933–9953.
- 115 H. Tae, E.-B. Kong and K. Park, *BMC Bioinformatics*, 2007, **8**, 327.
- 116 M. H. Li, P. M. Ung, J. Zajkowski, S. Garneau-Tsodikova and D. H. Sherman, *BMC Bioinformatics*, 2009, **10**, 1.
- 117 S. Anand and D. Mohanty, in *Handbook of Research on Computational and Systems Biology: Interdisciplinary Applications*, IGI Global snippet, USA: IGI G., 2011, pp. 414–439.
- 118 G. Yadav, R. S. Gokhale and D. Mohanty, *Nucleic Acids Res.*, 2003, **31**, 3654–3658.
- 119 M. Z. Ansari, G. Yadav, R. S. Gokhale and D. Mohanty, *Nucleic Acids Res.*, 2004, **32**, W405–W413.
- 120 P. Kamra, R. S. Gokhale and D. Mohanty, *Nucleic Acids Res.*, 2005, **33**, W220–W225.
- 121 B. O. Bachmann and J. Ravel, *Methods Enzymol.*, 2009, **458**, 181–217.
- 122 S. Anand, M. V. R. Prasad, G. Yadav, N. Kumar, J. Shehara, M. Z. Ansari and D. Mohanty, *Nucleic Acids Res.*, 2010, gkq340.
- 123 M. Röttig, M. H. Medema, K. Blin, T. Weber, C. Rausch and O. Kohlbacher, *Nucleic Acids Res.*, 2011, **39**, 362–367.
- 124 C. Prieto, *Characterization of nonribosomal peptide synthetases with NRPSp*, Springer P., 2016.
- 125 C. Prieto, C. Garcia-estrada, D. Lorenzana and J. F. Martin, *Bioinformatics*, 2012, **28**, 426–427.
- 126 Y. Minowa, M. Araki and M. Kanehisa, *J. Mol. Biol.*, 2007, **368**, 1500–1517.
- 127 Y. M. Chiang, E. Szewczyk, T. Nayak, A. D. Davidson, J. F. Sanchez, H. C. Lo, W. Y. Ho, H. Simityan, E. Kuo, A. Praseuth, K. Watanabe, B. R. Oakley and C. C. C. Wang, *Chem. Biol.*, 2008, **15**, 527–532.
- 128 Y. Ninomiya, K. Suzuki, C. Ishii and H. Inoue, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 12248–12253.
- 129 S. Krappmann, C. Sasse and G. H. Braus, *Eukaryot. Cell*, 2006, **5**, 212–215.
- 130 S. Pöggeler and U. Kück, *Gene*, 2006, **378**, 1–10.
- 131 M. E. Da Silva Ferreira, M. R. V. Z. Kress, M. Savoldi, M. H. S. Goldman, A. Härtl, T. Heinekamp, A. A. Brakhage and G. H. Goldman, *Eukaryot. Cell*, 2006, **5**, 207–211.
- 132 Y. M. Chiang, M. Ahuja, C. E. Oakley, R. Entwistle, A. Asokan, C. Zutz, C. C. C. Wang and B. R. Oakley, *Angew. Chemie - Int. Ed.*, 2016, **55**, 1662–1665.
- 133 Y. M. Chiang, E. Szewczyk, A. D. Davidson, N. Keller, B. R. Oakley and C. C. C. Wang, *J. Am. Chem. Soc.*, 2009, **131**, 2965–2970.
- 134 Y. M. Chiang, S. L. Chang, B. R. Oakley and C. C. C. Wang, *Curr. Opin. Chem. Biol.*, 2011, **15**, 137–143.
- 135 L. B. Pickens, Y. Tang and Y.-H. Chooi, *Annu. Rev. Chem. Biomol. Eng.*, 2011, **2**, 211.
- 136 Y. M. Chiang, C. E. Oakley, M. Ahuja, R. Entwistle, A. Schultz, S. L. Chang, C. T. Sung, C. C. C. Wang and B. R. Oakley, *J. Am. Chem. Soc.*, 2013, **135**, 7720–7731.
- 137 J. Yaegashi, B. R. Oakley and C. C. C. Wang, *J. Ind. Microbiol. Biotechnol.*, 2014, **41**, 433–442.
- 138 K. Sakai, H. Kinoshita, T. Shimizu and T. Nihira, *J. Biosci. Bioeng.*, 2008, **106**, 466–472.
- 139 W. B. Yin, Y. H. Chooi, A. R. Smith, R. A. Cacho, Y. Hu, T. C. White and Y. Tang, *ACS Synth. Biol.*, 2013, **2**, 629–634.
- 140 L. M. Petersen, D. K. Holm, C. H. Gotfredsen, U. H. Mortensen and T. O. Larsen, *ChemBioChem*, 2015, **16**, 2200–2204.
- 141 B. A. Pfeifer and C. Khosla, *Microbiol. Mol. Biol. Rev.*, 2001, **65**, 106–118.
- 142 K. Scherlach and C. Hertweck, *Org. Biomol. Chem.*, 2009, **7**, 1753–1760.
- 143 G. L. Challis, *Microbiology*, 2008, **154**, 1555–1569.
- 144 J. W. Bok, R. Ye, K. D. Clevenger, D. Mead, M. Wagner, A. Krerowicz, J. C. Albright, A. W. Goering, P. M. Thomas, N. L. Kelleher, N. P. Keller and C. C. Wu, *BMC Genomics*, 2015, **16**, 343.
- 145 Y. H. Chooi and P. S. Solomon, *Front. Microbiol.*, 2014, **5**, 640.
- 146 A. A. Brakhage, *Nat. Rev. Microbiol.*, 2013, **11**, 21–32.
- 147 P. K. Chang, K. C. Ehrlich, J. Yu, D. Bhatnagar and T. E. Cleveland, *Appl. Environ. Microbiol.*, 1995, **61**, 2372–2377.
- 148 F. Y. Lim, J. F. Sanchez, C. C. C. Wang and N. P. Keller, *Methods Enzymol.*, 2012, **517**, 303.
- 149 S. Bergmann, J. Schumann, K. Scherlach, C. Lange, A. A. Brakhage and C. Hertweck, *Nat. Chem. Biol.*, 2007, **3**, 213–217.
- 150 J. L. Sorensen, F. T. Hansen, T. E. Sondergaard, D. Staerk, T. V. Lee, R. Wimmer, L. G. Klitgaard, S. Purup, H. Giese and R. J. N. Frandsen, *Environ. Microbiol.*, 2012, **14**, 1159–1170.
- 151 Y. H. Chooi, J. Fang, H. Liu, S. G. Filler, P. Wang and Y. Tang, *Org. Lett.*, 2013, **15**, 780–783.

- 152 K. W. Von Bargaen, E. M. Niehaus, K. Bergander, R. Brun, B. Tudzynski and H. U. Humpf, *J. Nat. Prod.*, 2013, **76**, 2136–2140.
- 153 M. Ahuja, Y. M. Chiang, S. L. Chang, M. B. Praseuth, R. Entwistle, J. F. Sanchez, H. C. Lo, H. H. Yeh, B. R. Oakley and C. C. C. Wang, *J. Am. Chem. Soc.*, 2012, **134**, 8212–8221.
- 154 H. H. Yeh, Y. M. Chiang, R. Entwistle, M. Ahuja, K. H. Lee, K. S. Bruno, T. K. Wu, B. R. Oakley and C. C. C. Wang, *Appl. Microbiol. Biotechnol.*, 2012, **96**, 739–748.
- 155 J. W. Bok and N. P. Keller, *Eukaryot. Cell*, 2004, **3**, 527–535.
- 156 R. M. Perrin, N. D. Fedorova, W. B. Jin, R. A. Cramer, J. R. Wortman, H. S. Kim, W. C. Nierman and N. P. Keller, *PLoS Pathog.*, 2007, **3**, e50.
- 157 J. W. Bok, D. Hoffmeister, L. A. Maggio-Hall, R. Murillo, J. D. Glasner and N. P. Keller, *Chem. Biol.*, 2006, **13**, 31–37.
- 158 S. Dhingra, A. L. Lind, H. C. Lin, Y. Tang, A. Rokas and A. M. Calvo, *PLoS One*, 2013, **8**, e77147.
- 159 R. M. Duran, J. W. Cary and A. M. Calvo, *Appl. Microbiol. Biotechnol.*, 2007, **73**, 1158–1168.
- 160 A. N. Patananan, J. M. Palmer, G. S. Garvey, N. P. Keller and S. G. Clarke, *J. Biol. Chem.*, 2013, **288**, 14032–14045.
- 161 Y. Reyes-Dominguez, J. W. Bok, H. Berger, E. K. Shwab, A. Basheer, A. Gallmetzer, C. Scazzocchio, N. Keller and J. Strauss, *Mol. Microbiol.*, 2010, **76**, 1376–1386.
- 162 H. B. Bode, M. Walker and A. Zeeck, *European J. Org. Chem.*, 2000, **2000**, 1451–1456.
- 163 K. Ochi, *Biosci. Biotechnol. Biochem.*, 2007, **71**, 1373–1386.
- 164 S. M. Fang, C. J. Wu, C. W. Li and C. Bin Cui, *Mar. Drugs*, 2014, **12**, 1788–1814.
- 165 M. W. Xia, C. Bin Cui, C. W. Li and C. J. Wu, *Mar. Drugs*, 2014, **12**, 1545–1568.
- 166 M.-W. Xia, C.-B. Cui, C.-W. Li, C.-J. Wu, J.-X. Peng and D.-H. Li, *Mar. Drugs*, 2015, **13**, 5219–5236.
- 167 Y. J. Chai, C. Bin Cui, C. W. Li, C. J. Wu, C. K. Tian and W. Hua, *Mar. Drugs*, 2012, **10**, 559–582.
- 168 C. J. Wu, L. Yi, C. Bin Cui, C. W. Li, N. Wang and X. Han, *Mar. Drugs*, 2015, **13**, 2465–2487.
- 169 N. Wang, C. Cui and C.-W. Li, *Arch. Pharm. Res.*, 2016, **39**, 762–770.
- 170 Y. Dong, C. Bin Cui, C. W. Li, W. Hua, C. J. Wu, T. J. Zhu and Q. Q. Gu, *Mar. Drugs*, 2014, **12**, 4326–4352.
- 171 J. M. Jin, S. Lee, J. Lee, S. R. Baek, J. C. Kim, S. H. Yun, S. Y. Park, S. Kang and Y. W. Lee, *Mol. Microbiol.*, 2010, **76**, 456–466.
- 172 J. Sekiguchi and G. M. Gaucher, *Appl. Environ. Microbiol.*, 1977, **33**, 147–158.
- 173 M. S. Ramirez and M. E. Tolmasky, *Drug Resist. Updat.*, 2010, **13**, 151–171.
- 174 E. K. Shwab, J. W. Bok, M. Tribus, J. Galehr, S. Graessle and N. P. Keller, *Eukaryotic Cell*, 2007, **6**, 1656–1664.
- 175 H. W. Nützmann, J. Fischer, K. Scherlach, C. Hertweck and A. A. Brakhagea, *Appl. Environ. Microbiol.*, 2013, **79**, 6102–6109.
- 176 R. H. Cichewicz, *Nat. Prod. Rep.*, 2010, **27**, 11–22.
- 177 Y. M. Chung, C. K. Wei, D. W. Chuang, M. El-Shazly, C. T. Hsieh, T. Asai, Y. Oshima, T. J. Hsieh, T. L. Hwang, Y. C. Wu and F. R. Chang, *Bioorganic Med. Chem.*, 2013, **21**, 3866–3872.
- 178 R. B. Williams, J. C. Henrikson, A. R. Hoover, A. E. Lee and R. H. Cichewicz, *Org. Biomol. Chem.*, 2008, **6**, 1895–1897.
- 179 X. Wang, J. G. Sena Filho, A. R. Hoover, J. B. King, T. K. Ellis, D. R. Powell and R. H. Cichewicz, *J. Nat. Prod.*, 2010, **73**, 942–948.
- 180 J. W. Bok, Y.-M. Chiang, E. Szewczyk, Y. Reyes-Dominguez, A. D. Davidson, J. F. Sanchez, H.-C. Lo, K. Watanabe, J. Strauss, B. R. Oakley, C. C. C. Wang and N. P. Keller, *Nat. Chem. Biol.*, 2009, **5**, 462–464.
- 181 L. Studt, S. M. R. Rösler, I. Burkhardt, B. Arndt, M. Freitag, H.-U. Humpf, J. S. Dickschat and B. Tudzynski, *Environ. Microbiol.*, 2016, **18**, 4037–4054.
- 182 J. C. Henrikson, A. R. Hoover, P. M. Joyner and R. H. Cichewicz, *Org. Biomol. Chem.*, 2009, **7**, 435–438.
- 183 J. C. Albright, M. T. Henke, A. A. Soukup, R. A. McClure, R. J. Thomson, N. P. Keller and N. L. Kellerher, *ACS Chem. Biol.*, 2015, **10**, 1535–1541.
- 184 S. Bergmann, A. N. Funk, K. Scherlach, V. Schroeckh, E. Shelest, U. Horn, C. Hertweck and A. A. Brakhage, *Appl. Environ. Microbiol.*, 2010, **76**, 8143–8149.
- 185 M. R. Andersen, J. B. Nielsen, A. Klitgaard, L. M. Petersen, M. Zachariasen, T. J. Hansen, L. H. Blicher, C. H. Gotfredsen, T. O. Larsen, K. F. Nielsen and U. H. Mortensen, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, E99–E107.
- 186 N. P. Keller, G. Turner and J. W. Bennett, *Nat Rev Microbiol.*, 2005, **3**, 937–947.
- 187 S. V. S. Ipcho, J. K. Hane, E. A. Antoni, D. Ahren, B. Henrissat, T. L. Friesen, P. S. Solomon and R. P. Oliver, *Mol. Plant Pathol.*, 2012, **13**, 531–545.
- 188 Y.-H. Chooi, P. S. Solomon, J. M. Palmer and A. Trchounian, *Front. Microbiol.*, 2014, **19**, 640.
- 189 Z. Wang, M. Gerstein and M. Snyder, *Nat. Rev. Genet.*, 2009, **10**, 57–63.
- 190 X.-W. Zhang, L.-J. Jia, Y. Zhang, G. Jiang, X. Li, D. Zhang and W.-H. Tang, *Plant Cell*, 2012, **24**, 5159–5176.
- 191 S. Bertrand, N. Bohni, S. Schnee, O. Schumpp, K. Gindro and J. L. Wolfender, *Biotechnol. Adv.*, 2014, **32**, 1180–1204.
- 192 W. Yin and N. P. Keller, *J. Microbiol.*, 2011, **49**, 329–339.
- 193 H. Nützmann, Y. Reyes-dominguez, K. Scherlach, V. Schroeckh and F. Horn, *PNAS*, 2011, **108**, 14282–14287.
- 194 A. Marmann, A. H. Aly, W. Lin, B. Wang and P. Proksch, *Mar. Drugs*, 2014, **12**, 1043–1065.
- 195 K. Nonaka, T. Abe, M. Iwatsuki, M. Mori, T. Yamamoto, K. Shiomi, S. Mura and R. Masuma, *J. Antibiot. (Tokyo)*, 2011, **64**, 769–77491.
- 196 J. P. Wang, W. Lin, V. Wray, D. Lai and P. Proksch, *Tetrahedron Lett.*, 2013, **54**, 2492–2496.
- 197 K. M. Zuck, S. Shipley and D. J. Newman, *J. Nat. Prod.*, 2011, **74**, 1653–1657.
- 198 T. Degenkolb, S. Heinze, B. Schlegel, G. Strobel and U. Grafe, *Biosci. Biotechnol. Biochem.*, 2002, **66**, 883–886.
- 199 A. R. B. Ola, D. Thomy, D. Lai, H. Brötz-Oesterhelt and P.

- Proksch, *J. Nat. Prod.*, 2013, **76**, 2094–2099.
- 200 C. C. König, K. Scherlach, V. Schroeckh, F. Horn, S. Nietzsche, A. A. Brakhage and C. Hertweck, *ChemBioChem*, 2013, **14**, 938–942.
- 201 R. Chávez, F. Fierro, R. O. García-Rico and I. Vaca, *Front. Microbiol.*, 2015, **6**.
- 202 F. Zhu, G. Chen, X. Chen, M. Huang and X. Wan, *Chem.*, 2011, **47**, 767.
- 203 C. Li, J. Zhang, C. Shao, W. Ding, Z. She and Y. Lin, *Chem. Nat. Compd.*, 2011, **47**, 382.
- 204 D. C. Oh, P. R. Jensen, C. A. Kauffman and W. Fenical, *Bioorganic Med. Chem.*, 2005, **13**, 5267–5273.
- 205 M. Cueto, P. R. Jensen, C. Kauffman, W. Fenical, E. Lobkovsky and J. Clardy, *J. Nat. Prod.*, 2001, **64**, 1444–1446.
- 206 D. C. Oh, C. A. Kauffman, P. R. Jensen and W. Fenical, *J. Nat. Prod.*, 2007, **70**, 515–520.
- 207 H. B. Park, H. C. Kwon, C. H. Lee and H. O. Yang, *J. Nat. Prod.*, 2009, **72**, 248–252.
- 208 P. A. Paranagama, E. M. K. Wijeratne and A. A. L. Gunatilaka, *J. Nat. Prod.*, 2007, **70**, 1939–1945.
- 209 M. E. Rateb and R. Ebel, *Nat. Prod. Rep.*, 2011, **28**, 290–344.
- 210 R. C. Bonugli-Santos, M. R. dos S. Vasconcelos, M. R. Z. Passarini, G. A. L. Vieira, V. C. P. Lopes, P. H. Mainardi, J. A. dos Santos, L. de A. Duarte, I. V. R. Otero, A. M. da S. Yoshida, V. A. Feitosa, A. Pessoa and L. D. Sette, *Front. Microbiol.*, 2015, **6**, 269.
- 211 K. Li, Y.-W. Chung-Davidson, U. Bussy and W. Li, *Mar. Drugs*, 2015, **13**, 2694–2713.
- 212 C. Christophersen, O. Crescente, J. C. Frisvad, L. Gram, J. Nielsen, H. Nielsen and L. Rahbaek, *Mycopathologia*, 1999, **143**, 135–138.
- 213 Y. Wang, Z. Lu, K. Sun and W. Zhu, *Mar. Drugs*, 2011, **9**, 535–542.
- 214 Y. Wang, J. Zheng, P. Liu, W. Wang and W. Zhu, *Mar. Drugs*, 2011, **9**, 1368–1378.
- 215 J. Zheng, Z. Xu, Y. Wang, K. Hong, P. Liu and W. Zhu, *J. Nat. Prod.*, 2010, **73**, 1133–1137.
- 216 H. B. Bode, M. Walker and A. Zeeck, *European J. Org. Chem.*, 2000, **2000**, 3185–3193.
- 217 K. Scherlach, J. Schuemann, H.-M. Dahse and C. Hertweck, *J. Antibiot. (Tokyo)*, 2010, **63**, 375.
- 218 G. W. Bruehl, R. L. Millar and B. Cunfer, *Can. J. Plant Sci.*, 1969, **49**, 235–246.
- 219 D. P. Overy, J. Smedsgaard, J. C. Frisvad, R. K. Phipps and U. Thrane, *J. Appl. Microbiol.*, 2006, **101**, 1292–1300.
- 220 K. Scherlach and C. Hertweck, *Org. Biomol. Chem.*, 2006, **4**, 3517–3520.
- 221 A. Sarkar, A. N. Funk, K. Scherlach, F. Horn, V. Schroeckh, P. Chankhamjon, M. Westermann, M. Roth, A. A. Brakhage, C. Hertweck and U. Horn, *J. Biotechnol.*, 2012, **160**, 64–71.
- 222 K. Scherlach, A. Sarkar, V. Schroeckh, H. M. Dahse, M. Roth, A. A. Brakhage, U. Horn and C. Hertweck, *ChemBioChem*, 2011, **12**, 1836–1839.
- 223 M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna and E. Charpentier, *Science (80- )*, 2012, **337**, 816–822.
- 224 J. A. Doudna and E. Charpentier, *Science (80- )*, 2014, **346**, 1258096–1258096.
- 225 C. S. Nodvig, J. B. Nielsen, M. E. Kogle and U. H. Mortensen, *PLoS One*, 2015, **10**, 1–18.
- 226 C. Zhang, X. Meng, X. Wei and L. Lu, *Fungal Genet. Biol.*, 2016, **86**, 47–57.
- 227 T. Arazoe, K. Miyoshi, T. Yamato, T. Ogawa, S. Ohsato, T. Arie and S. Kuwata, *Biotechnol. Bioeng.*, 2015, **112**, 2543–2549.
- 228 J. Weber, V. Valiante, C. S. Nødvig, D. J. Mattern, R. A. Slotkowski, U. H. Mortensen and A. A. Brakhage, *ACS Synth. Biol.*, 2017, **6**, 62–68.
- 229 V. Meyer, M. R. Andersen, A. A. Brakhage, G. H. Braus, M. X. Caddick, T. C. Cairns, R. P. de Vries, T. Haarmann, K. Hansen, C. Hertz-Fowler, S. Krappmann, U. H. Mortensen, M. A. Peñalva, A. F. J. Ram and R. M. Head, *Fungal Biol. Biotechnol.*, 2016, **3**, 6.
- 230 M. L. Nielsen, T. Isbrandt, K. B. Rasmussen, U. Thrane, J. B. Hoof, T. O. Larsen and U. H. Mortensen, *PLoS One*, 2017, **12**, e0169712.
- 231 N. Bohni, M. L. Cordero-Maldonado, J. Maes, D. Siverio-Mota, L. Marcourt, S. Munck, A. R. Kamuhabwa, M. J. Moshi, C. V. Esguerra, P. A. M. de Witte, A. D. Crawford and J. L. Wolfender, *PLoS One*, 2013, **8**, e64006.
- 232 M. Lambert, J. L. Wolfender, D. Stærk, S. B. Christensen, K. Hostettmann and J. W. Jaroszewski, *Anal. Chem.*, 2007, **79**, 727–735.
- 233 S. Bertrand, O. Schumpp, N. Bohni, A. Bujard, A. Azzollini, M. Monod, K. Gindro and J. L. Wolfender, *J. Chromatogr. A*, 2013, **1292**, 219–228.
- 234 G. Glauser, K. Gindro, J. Fringeli, J. P. De Joffrey, S. Rudaz and J.-L. Wolfender, *J. Agric. Food Chem.*, 2009, **57**, 1127–1134.
- 235 B. C. Covington, J. A. McLean and B. O. Bachmann, *Nat. Prod. Rep.*, 2016, **34**, 6–24.
- 236 J. C. Nielsen, S. Grijseels, S. Prigent, B. Ji, J. Dainat, K. F. Nielsen, J. C. Frisvad, M. Workman and J. Nielsen, *Nat. Microbiol.*, 2017, **2**, 17044.
- 237 P. Wiemann, C. M. K. Sieber, K. W. von Bargen, L. Studt, E. M. Niehaus, J. J. Espino, K. Huß, C. B. Michielse, S. Albermann, D. Wagner, S. V. Bergner, L. R. Connolly, A. Fischer, G. Reuter, K. Kleigrewe, T. Bald, B. D. Wingfield, R. Ophir, S. Freeman, M. Hippler, K. M. Smith, D. W. Brown, R. H. Proctor, M. Münsterkötter, M. Freitag, H. U. Humpf, U. Güldener and B. Tudzynski, *PLoS Pathog.*, 2013, **9**, e1003475.
- 238 S. Bertrand, O. Schumpp, N. Bohni, M. Monod, K. Gindro and J. L. Wolfender, *J. Nat. Prod.*, 2013, **76**, 1157–1165.
- 239 W. J. Moree, V. V. Phelan, C. H. Wu, N. Bandeira, D. S. Cornett, B. M. Duggan and P. C. Dorrestein, *Proc Natl Acad Sci U S A*, 2012, **109**, 13811–13816.



## 4. DES ESPÈCES FONGIQUES TOXINOGENES

### 4.1. *P. verrucosum* et *P. nordicum*

#### 4.1.1 Le genre *Penicillium*

*Penicillium* est un genre de champignon filamenteux de l'ancien groupe des deutéromycètes (champignons à forme imparfaite) bien que quelques espèces présentent des formes parfaites (appareil reproducteur sexué) et soient alors classées dans le genre *Eupenicillium*. Ce genre comprend entre 100 et 300 espèces de la division *Ascomycota*, classe des *Eurotiomycetes*, sous-classe des *Eurotiomycetidae*, ordre des *Eurotiales* et famille des *Tichocomaceae*. Le genre *Penicillium* est divisé en quatre sous-genres : *Aspergilloïdes*, *Penicillium*, *Biverticillium* et *Furcatum*<sup>7</sup>. Ces moisissures sont retrouvées dans le sol, les végétaux en décomposition, le bois, les produits alimentaires secs ainsi que les fruits et légumes, mais également sur des matériaux de construction dans des environnements endommagés par l'eau ou dans les poussières domestiques<sup>8</sup>.

Décrit pour la première fois en 1809, ce genre de moisissure fut nommé à partir de la racine latine exprimant la forme de pinceau (ou pénéicille) selon laquelle les chaînes de ses conidies sont organisées (Figure 2). Les *Penicillium* sont reconnaissables par leur conidiophores dressés qui peuvent être ramifiés et se terminent par des phialides. Ces dernières sont disposées en verticilles à l'extrémité des conidiophores (Figure 3). Le nombre de ramifications entre le conidiophore et les spores permet leur classification parmi les monoverticillés, les biverticillés, les terverticillés ou les tétraverticillés (Figure 2).

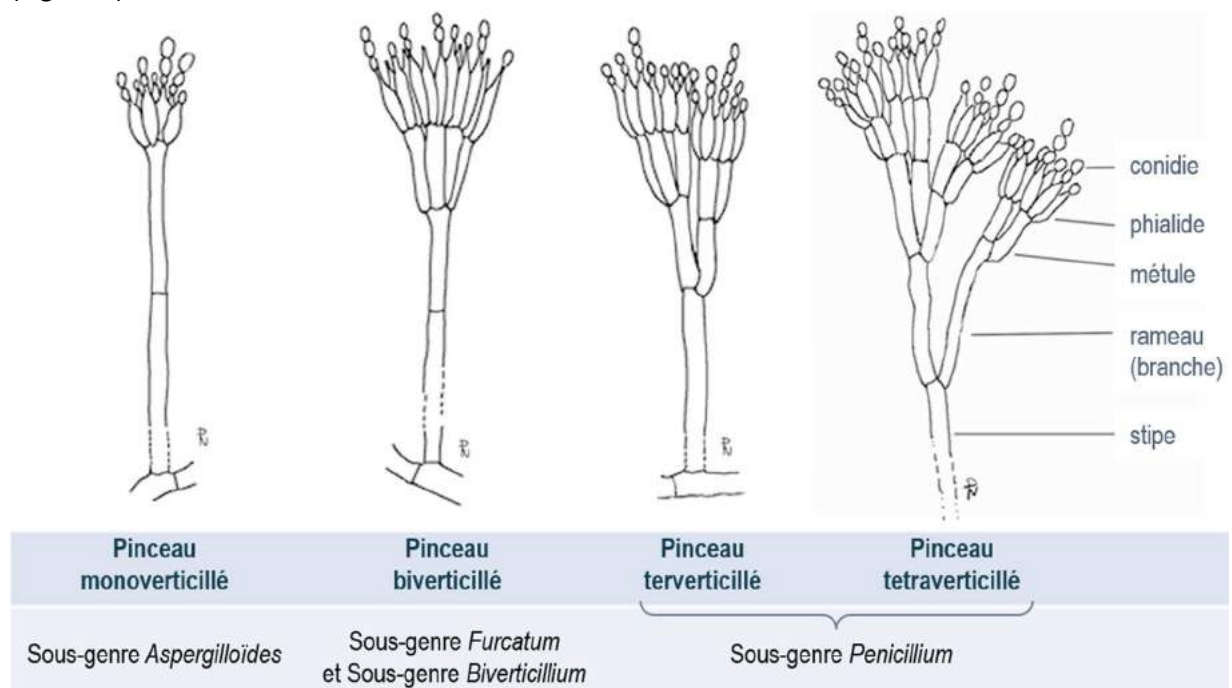


Figure 2. Classification des *Penicillium* selon le degré de ramification de leur mycélium<sup>9</sup>.

Le genre *Penicillium* est connu en particulier pour la production de pénicilline par l'une de ses espèces. Ce premier antibiotique, découvert par sérendipité, a ouvert le champ du développement de médicaments d'origine fongique. De nombreux médicaments ont ainsi été développés à partir de métabolites secondaires de *Penicillium*, tels que des agents hypocholestérolémiants (compactine de *P. citrinum*) ou des immunosupresseurs (acide mycophénolique sécrété par *P. brevicompactum*)<sup>10</sup>.



Outre leurs vertus thérapeutiques, les métabolites de certains *Penicillium* sont utilisés pour leurs propriétés organoleptiques, notamment pour la production de fromages tels que le roquefort (*P. roqueforti*) et le camembert (*P. camemberti*). Ces différentes propriétés suscitent un intérêt pour l'étude du métabolisme secondaire de ces moisissures dans divers domaines industriels.

Néanmoins, le genre *Penicillium* est également connu pour sa production de mycotoxines, associées pour la plupart à la détérioration de denrées alimentaires. Parmi les mycotoxines les mieux étudiées de *Penicillium*, citons l'acide pénicillique, la citrinine, des mycotoxines trémorgènes (le pénitrem, le territrem, le verruculogène), les ochratoxines A et B, la patuline, la roquefortine C, la verrucosidine, la viomelleine et la viridicatine<sup>8</sup>. Outre les effets irritants et allergiques qu'elles sont susceptibles de déclencher lorsqu'elles sont émises dans les poussières domestiques, ces toxines sont à l'origine d'intoxications via l'alimentation humaine ou animale. C'est pourquoi il est important de pouvoir contrôler et maîtriser la contamination des denrées alimentaires.

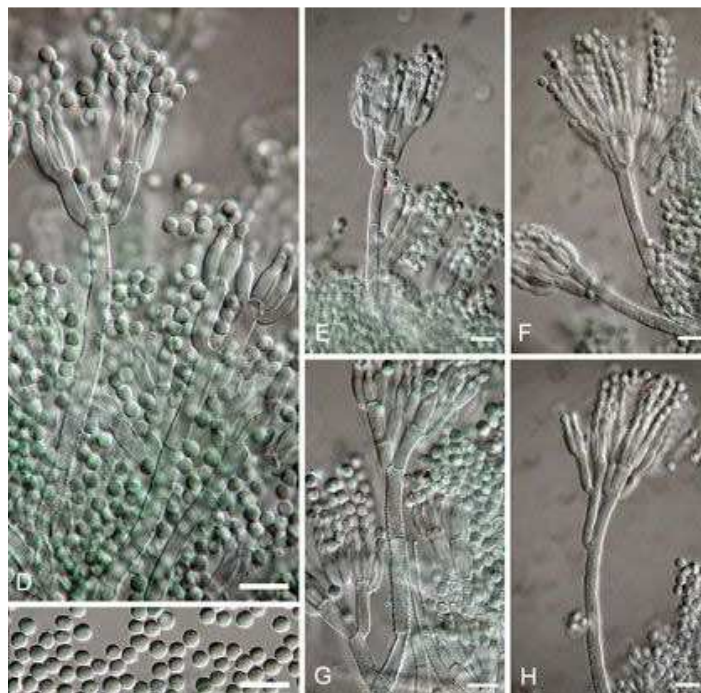


Figure 3. *Penicillium verrucosum*. Conidiophores (D-H). Conidies (I). Barre d'échelle =10  $\mu$ m. Source : Frisvad et al. 2004<sup>11</sup>

### 4.1.2. Propriétés de *P. verrucosum* et *P. nordicum*

Les deux espèces de *Penicillium* concernées par cette étude, *P. verrucosum* et *P. nordicum*, sont deux espèces de *Penicillium* terverticillés issus du sous-genre *Penicillium*, de la section *Viridicata* et de la série *Verrucosa*<sup>11</sup>. *P. verrucosum* et *P. nordicum* sont deux espèces capables de croître dans des conditions de faible activité en eau ( $a_w$ ) (jusqu'à moins de 0.80)<sup>12</sup>. Cette propriété leur permet de se développer dans les stocks alimentaires. En raison d'une croissance uniquement à basse température (de 0 à 30°C, avec une température optimale de 20°C), ces deux espèces ont un habitat qui s'étend sur les zones tempérées (Europe, USA, Russie et Canada principalement). La destruction de ces denrées, ainsi que leur consommation par l'Homme et/ou les animaux, placent ces champignons filamenteux au cœur d'importantes problématiques économiques et sanitaires.

*P. verrucosum* est un parasite saprophyte qui infeste les céréales pendant leur stockage, pouvant les rendre ainsi impropres à la consommation humaine. Dans certaines régions d'Europe, cette espèce est détectée sur plus de 80% des échantillons de blés récoltés<sup>13</sup>. En raison de sa forte prévalence et de son fort pouvoir pathogène, *P. verrucosum* figure parmi les espèces de champignons filamenteux toxigènes associés à la nourriture les mieux connus<sup>8</sup>. Bien qu'occasionnellement détectés sur des céréales, *P. nordicum* présente quant à lui une affinité particulière pour les substrats hyper-protéinés tels que les viandes séchées, les poissons salés ou les fromages<sup>11</sup>. *P. nordicum* est systématiquement isolé de produits conservés sous réfrigération pendant leur fabrication<sup>12</sup>.

#### 4.1.3. Critères de classification de *P. verrucosum* et *P. nordicum*

Avant d'être distingués, *P. verrucosum* et *P. nordicum* étaient groupés de manière indifférenciée sous le genre *P. viridicatum*<sup>14</sup>. Dans un second temps, un consensus fut atteint parmi les taxonomistes qui convinrent que tous les *Penicillium* producteurs d'ochratoxine A seraient réunis sous une même espèce nommée alors *P. verrucosum*<sup>15</sup>. Plus tard, cette classification fut scindée en deux espèces distinctes d'après des critères qui seront précisés dans ce chapitre : *P. verrucosum* et *P. nordicum*<sup>16</sup>.

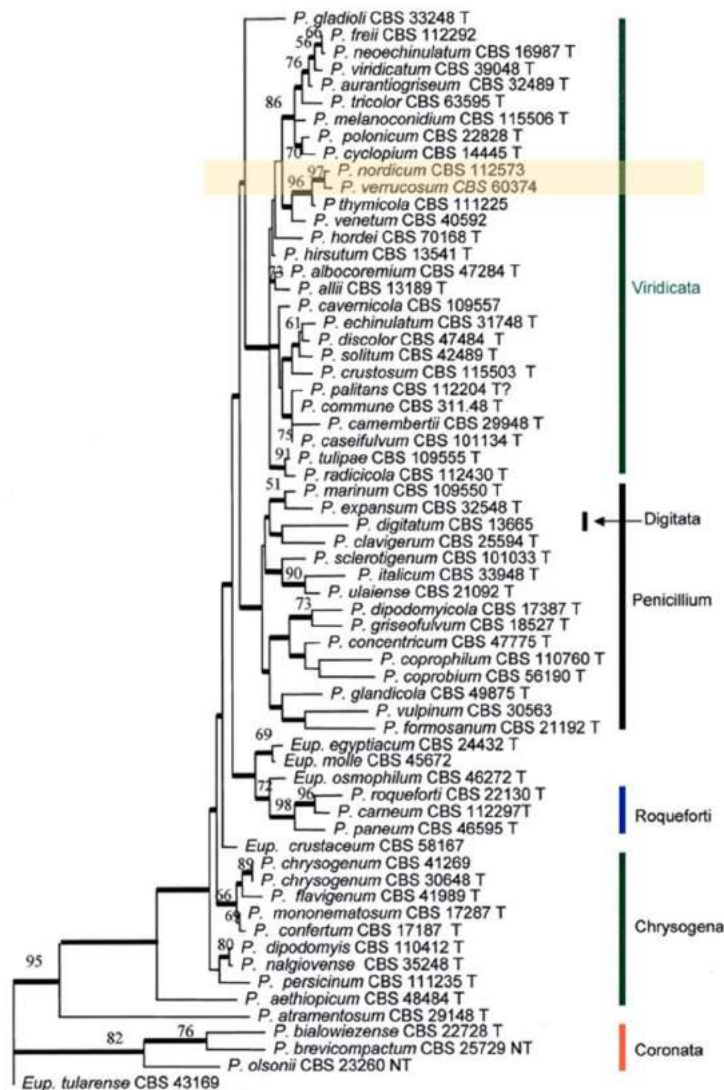


Figure 4. Proximité phylogénique de *P. verrucosum* et *P. nordicum* au sein des plus de 250 espèces référencées du genre *Penicillium*. Portion de l'arbre phylogénique du sous-genre *Penicillium* selon le gène de la  $\beta$ -tubuline<sup>17</sup>.

Cette précision historique permet d'illustrer l'étroite proximité taxonomique des deux espèces étudiées (Figure 4) et introduit la problématique liée à la caractérisation des moisissures dans les études à travers les époques. De nombreux articles évoquent les problématiques engendrées par les variations de classifications phylogéniques<sup>11,18</sup>. Normalement, les métabolites suivent leurs souches productrices au fil des modifications taxonomiques en lien avec la nouvelle dénomination de celles-ci. Néanmoins, de nombreux métabolites n'ont pas pu être réappropriés en raison d'un manque de précisions quant à la souche étudiée à l'origine. Ainsi, certains métabolites secondaires de *P. verrucosum* et de *P. nordicum* peuvent être encore référencés en tant que métabolites de *P. viridicatum*. Dans ce contexte, et afin de permettre un suivi rigoureux des métabolites à travers les découvertes scientifiques, il est nécessaire de détailler les critères de caractérisation des souches analysées dans chaque étude.

La caractérisation de *P. verrucosum* et *P. nordicum* peut être évaluée selon des critères morphologiques, moléculaires et métaboliques<sup>11</sup>. Sur le plan morphologique, ces deux moisissures présentent de très nombreuses similitudes. Si elles sont aisément reconnaissables au sein des *Penicillium*, il est néanmoins délicat de les différencier microscopiquement puisque les deux espèces présentent les mêmes pinceaux asymétriques terverticillés ainsi que des conidiophores rugueux avec spores globuleuses lisses. À l'échelle macroscopique, l'unique facteur déterminant de la discrimination entre *P. verrucosum* et *P. nordicum* est le test de culture sur milieu YES agar (Yeast Extract Sucrose : Yeast extract (Difco) 2.0%, Sucrose 15.0%, MgSO<sub>4</sub>·7H<sub>2</sub>O 0.05%, CuSO<sub>4</sub>·5H<sub>2</sub>O 0.0005%, ZnSO<sub>4</sub>·7H<sub>2</sub>O 0.001%, Agar 2.0%, Distilled water)<sup>11</sup>. Alors que le revers de la culture présente une couleur blanc crème pour *P. nordicum*, il apparaît rouge-brun pour *P. verrucosum* (Figure 5). Néanmoins, Larsen *et al.* ont décrit deux souches de *P. verrucosum* ne présentant pas de revers brun lors de leur culture sur milieu YES agar<sup>16</sup>. Compte tenu de leur métabolome secondaire, ces deux champignons ont néanmoins été caractérisés comme *P. verrucosum* (production d'ochratoxine A, de verrucines et de verrucolones. Voir troisième critère de classification).

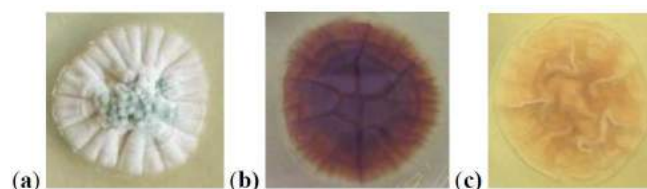


Figure 5. Culture sur milieu YES agar de *P. verrucosum* et *P. nordicum* après 7 jours à 25°C<sup>18</sup>. (a) aspect d'une colonie de *P. verrucosum* et *P. nordicum*. (b) revers de la colonie de *P. verrucosum*. (c) revers de la colonie de *P. nordicum*.

Le deuxième critère de classification de *P. verrucosum* et *P. nordicum* peut être investigué au niveau moléculaire par l'analyse de séquences génétiques conservées au cours de l'évolution. En effet, la séquence de la  $\beta$ -tubuline a (BenA) est un marqueur pertinent pour la caractérisation d'espèces fongiques et peut être utilisée en particulier pour l'étude des souches du genre *Penicillium*<sup>17</sup>. Pour ce faire, des analyses moléculaires permettent d'isoler et d'amplifier la séquence correspondante chez la souche étudiée et de la comparer par alignement de séquences avec la séquence standard répertoriée dans la littérature. L'utilisation de la séquence BenA s'explique d'une part par la grande conservation des séquences des exons permettant une amplification efficace de ces séquences d'ADN, et d'autre part par la variabilité au sein des introns qui a été démontrée comme représentative des différentes souches fongiques. Quoique moins pratiquée pour l'étude des *Penicillium*, la comparaison des

séquences ITS (les espaces inter-gènes codant pour les différentes sous-unités de l'ARN ribosomique) des souches étudiées avec des séquences standard peut également être réalisée pour la caractérisation d'espèces fongiques.

Le troisième critère d'identification de ces champignons est métabolique. Évoquée dans un article pionnier de Frisvad *et al.*<sup>19</sup> en 1983, la classification des *Penicillium* selon leurs métabolomes secondaires a permis de définir *P. verrucosum* et *P. nordicum* comme les deux uniques espèces de *Penicillium* productrices d'ochratoxine A<sup>16</sup>. La classification métabolique permet également de distinguer ces deux moisissures. Ainsi, Frisvad *et al.* définissent *P. verrucosum* comme étant la seule espèce de *Penicillium* produisant à la fois l'ochratoxine A et la citrinine<sup>11</sup>. L'observation de souches non productrices de citrinine a permis quant à elle de définir *P. nordicum* comme l'espèce de *Penicillium* produisant l'ochratoxine A en l'absence de citrinine. La sécrétion d'anacine est également définie comme une caractéristique propre à *P. nordicum*<sup>17</sup>. Les critères de classification sont résumés dans le Tableau 1.

		<i>P. verrucosum</i>	<i>P. nordicum</i>
<b>Références des souches</b>		CBS115508	CBS 112573
<b>Niches environnementales</b>		Céréales (blé principalement)	Substrat protéiné (viande, poisson, fromage)
<b>Caractérisation morphologique</b>	<b>Microscopique</b>	Pinceaux asymétriques terverticillés Conidiophores rugueux Spores globuleuses lisses	
	<b>Macroscopique</b>	Revers rouge-brun sur YES agar	Revers blanc crème sur YES agar
<b>Caractérisation moléculaire</b>	<b>BenA</b>	KF499576	AY674317
	<b>ITS</b>	NR_119495	
<b>Caractérisation métabolique</b>	Ochratoxine A		
		Production plus faible	Production plus importante
		Citrinine	Absence de citrinine
			Anacine

Tableau 1. Résumé des critères de classification de *P. verrucosum* et *P. nordicum*<sup>11</sup>.

#### 4.1.4. Métabolomes secondaires de *P. verrucosum* et *P. nordicum*

L'ochratoxine A et l'ochratoxine B sont deux mycotoxines connues pour être produites par *P. verrucosum* et *P. nordicum*<sup>20</sup>. Comme pour la plupart des mycotoxines, il n'existe pas de procédé de décontamination des céréales infestées. La stabilité chimique, thermique et biologique de la majeure partie des mycotoxines empêche leur destruction lors de la transformation des aliments. Ces métabolites sont donc fréquemment détectés dans les céréales, le vin, la bière et la viande de porc (essentiellement les abats) qui sont les principales sources d'exposition des humains aux ochratoxines dans les zones tempérées (en Europe et au Moyen-Orient). Ces molécules pourraient être associées à la néphropathie endémique des Balkans, une pathologie mortelle caractérisée par des dysfonctionnements rénaux qui ont touché les populations d'Europe de l'est<sup>21</sup>. Outre un effet néphrotoxique, ces composés sont responsables de lésions du tractus gastro-intestinal, du foie, du cerveau, du cœur et des tissus lymphatiques chez différents animaux. Il a également été rapporté des effets embryotoxiques, tératogènes et immunosuppresseurs à des doses supérieures aux doses néphrotoxiques. Enfin, l'ochratoxine A est classée parmi les molécules du groupe 2B par l'IARC, c'est-à-dire comme carcinogène potentiel pour l'Homme<sup>22</sup>. L'atome de chlore de l'ochratoxine A contribue fortement à sa toxicité. C'est pourquoi l'ochratoxine B, qui est le dérivé non chloré de l'ochratoxine A,

présente une toxicité au moins dix fois inférieure. D'autres molécules sont référencées en tant que métabolites secondaires produits à la fois par des souches de *P. verrucosum* et de *P. nordicum* : les verrucosapyrones A et B, le 2-méthylisobornéol, les verrucolones (la verrucolone et PC-2), LL-P880γ<sup>20,23</sup> (Figure 6).

Une vingtaine de composés ont quant à eux été spécifiquement observés dans des extraits de *P. verrucosum*, mais seuls certains d'entre eux sont communément considérés comme des métabolites secondaires de ce champignon. On trouve notamment la citrinine observée chez 87% des *P. verrucosum* infectant les céréales<sup>23</sup>. Cette molécule aux propriétés antibiotiques est elle aussi à l'origine d'une néphrotoxicité via la destruction de l'épithélium des tubules rénaux. Cette mycotoxine a également été associée à des propriétés hépatotoxiques, génotoxiques et carcinogènes<sup>24</sup>. Malgré ces effets, la teneur en citrinine dans l'alimentation n'est pas réglementée. Les autres métabolites référencés chez *P. verrucosum* sont les verrucines A, B et F, le 2-méthylisobornéol<sup>11,25</sup>, l'acide cyclopiazonique<sup>26</sup>, ainsi qu'une anthraquinone pouvant expliquer la couleur brune du revers des cultures de *P. verrucosum* sur YES agar<sup>23</sup> (Figure 6).

Outre les métabolites secondaires communs aux deux espèces de *Penicillium* étudiées, *P. nordicum* peut produire de l'anacine. Ce métabolite secondaire n'ayant jamais été observé chez *P. verrucosum*, il est considéré comme métabolite discriminant lors de la caractérisation de ces deux espèces. De plus, *P. nordicum* produit la sclérotigénine, la lumpidine et l'acide viridique<sup>11</sup> (Figure 6). Larsen *et al.* évoquent également la synthèse d'un métabolite nommé Met I par *P. nordicum*, ainsi que la détection par UV de quatre composés de la famille de l'anacine<sup>16</sup>. Une analyse par LC-MS a indiqué qu'il pourrait s'agir de deux isomères protonés de l'anacine ainsi que deux isomères à la fois protonés et hydroxylés (l'un d'eux étant nommé anacinol). Certains articles proposent une sous-classification des *P. nordicum*, l'un d'eux étant caractérisé par la synthèse de verrucolones, d'anacines et d'ochratoxine A, et par l'absence de production de sclérotigénine<sup>16</sup>.

Bien que la plupart des souches étudiées dans la littérature respectent ces observations quant à la sécrétion de métabolites secondaires spécifiques à l'une ou l'autre des espèces, certaines exceptions ont été soulignées. Notamment, lors de la caractérisation de 48 souches de *P. verrucosum* et *P. nordicum*, Larsen *et al.* ont détecté la production de verrucines par une souche de *P. nordicum*<sup>16</sup>.

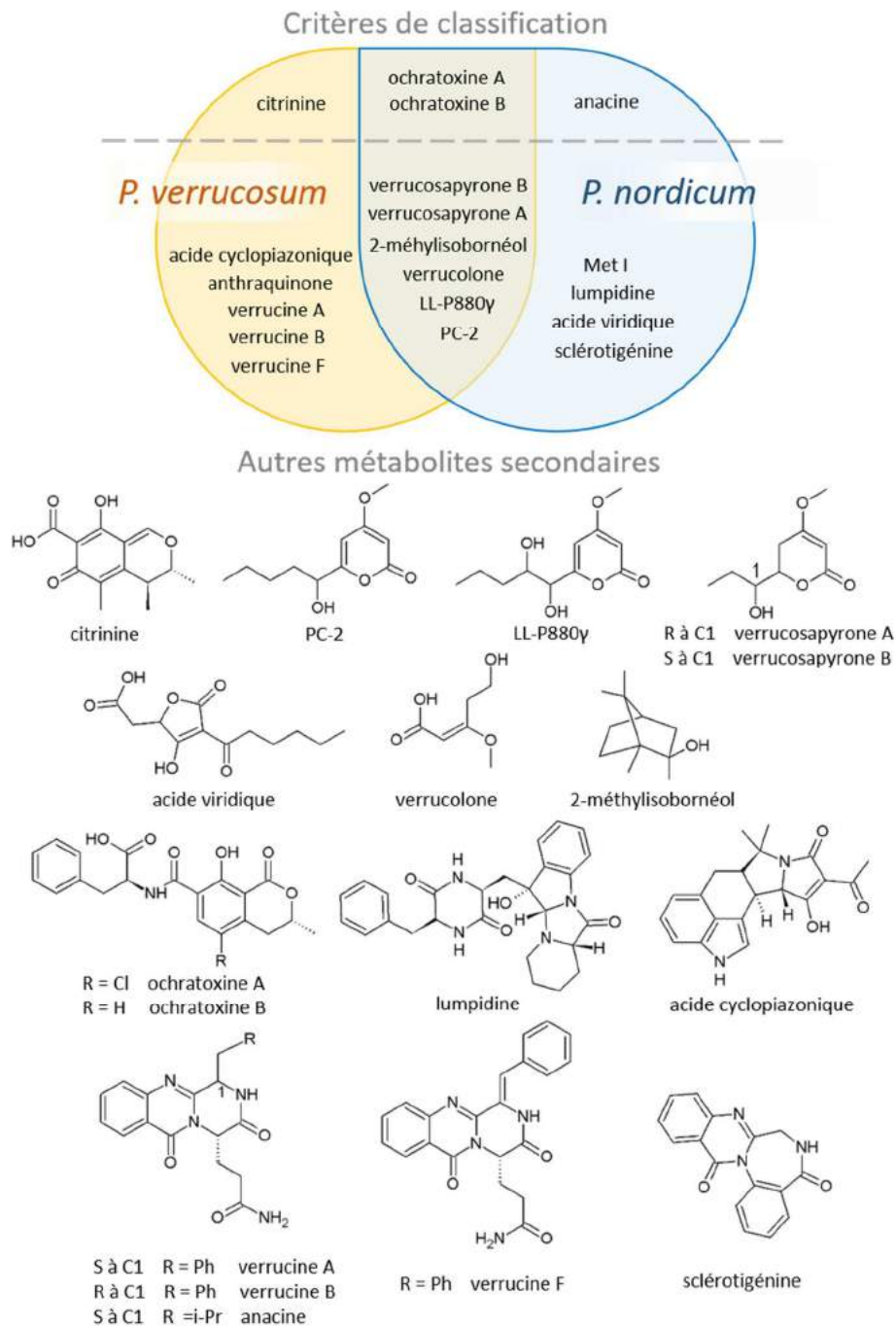


Figure 6. Métabolites secondaires de *P. verrucosum* et *P. nordicum* recensés dans la littérature

## 4.2. *Fusarium graminearum*, un champignon phytopathogène prédominant en Europe

### 4.2.1. Présentation de *F. graminearum*

Le genre *Fusarium* est composé de champignons filamenteux du groupe des ascomycètes, de la classe des *Sordariomycètes*, de l'ordre des *Hypocreales* et de la famille des *Nectriaceae*. Son nom provient de la forme en fuseau (du latin *fusus*) caractéristique de ses spores. En particulier, l'espèce *F. graminearum* est aisément identifiable grâce à ses macroconidies en forme de fins fuseaux incurvés et divisés en cinq ou six septa (Figure 7). Sur le plan macroscopique, *F. graminearum* présente un

mycélium dense et blanc. La pigmentation rouge visible sur le revers de la gélose vire au jaune lors de l'acidification du milieu.

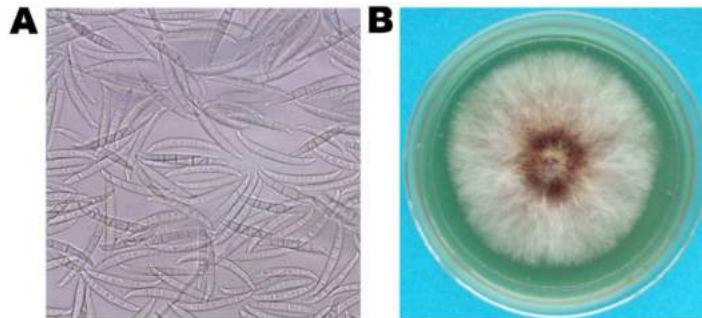


Figure 7. Observation macroscopique et microscopique de *F. graminearum*.

Les espèces du genre *Fusarium* comptent parmi les espèces fongiques les plus productives de métabolites secondaires. Certains d'entre eux présentent des propriétés d'intérêts industriels, tels que l'acide gibbérellique commercialisé en tant que promoteur de la croissance végétale<sup>27</sup>. À l'opposé, de nombreux métabolites secondaires toxiques synthétisés par les champignons du genre *Fusarium* sont responsables de pathologies chez les plantes, les animaux et l'Homme, que l'on appelle de manière générale fusarioses.

En particulier, *F. graminearum* est l'un des plus grands pathogènes des céréales et a été classé en quatrième position parmi les 10 champignons les plus phytopathogènes. Parce qu'il diminue la qualité des pieds infestés et peut les rendre impropres à la consommation, ce champignon est à l'origine d'importantes pertes économiques, particulièrement dans les régions tempérées du globe. L'infestation des champs par les spores peut intervenir au niveau des épis ou au niveau des racines lorsque les conditions climatiques sont favorables à son développement, c'est-à-dire au moment de la floraison. Les lésions causées par des insectes ou les ouvertures naturelles de la plante permettent alors la pénétration du champignon. La propagation du champignon et la sécrétion de mycotoxines sont ensuite responsables de la nécrose des tissus végétaux et de la perte de la qualité des grains. La fusariose de l'épi, ou *Fusarium Head Blight*, est l'une des maladies les plus fréquentes chez le blé et dont la décoloration des grains est un signe caractéristique (Figure 8).



Figure 8. Epi de blé infesté par *F. graminearum*.

#### 4.2.2. Le métabolome secondaire de *F. graminearum*

Certains des métabolites produits par *F. graminearum* font l'objet d'attentions particulières en raison de leur toxicité importante et leur teneur dans l'alimentation est réglementée (Figure 9). Notamment, les trichothécènes de type B (le déoxynivalénol (DON), le nivalénol, et leurs dérivés acétylés) sont suspectés d'être à l'origine de la virulence de ce champignon<sup>28</sup>. Lorsqu'il est ingéré, le nivalénol agit sur le système nerveux central et entraîne une diminution de l'appétit. Des retards de croissance, des néphropathies et des effets immunotoxiques et hématotoxiques sont également observés<sup>29</sup>. Le DON est quant à lui l'une des mycotoxines les plus fréquemment détectées sur les céréales, ce qui fait de lui l'un des produits fongiques les plus étudiés bien qu'il ne soit pas le plus toxiques des trichothécènes. Outre ses effets toxiques sur les plantes, le DON est responsable de pertes de poids, de retards de croissance, d'irritations intestinales et de la mort chez les animaux étudiés<sup>29</sup>. Cette toxine est également appelée vomitoxine en raison de son effet émétique via une action sur le système nerveux central lors d'intoxications aiguës. Les dérivés acétylés de ces deux mycotoxines, que sont le 3-acetyldeoxynivalénol, le 15-acetyldeoxynivalénol et le 4-acétylnivalénol (ou fusarénone X), sont souvent détectés dans les échantillons de céréales contaminés par le DON<sup>29</sup>. Dans la base de données AntiBase, d'autres métabolites secondaires précurseurs ou dérivés du DON sont référencés comme étant produits par *F. graminearum* ou *Gibberella zeae* (la forme sexuée de *F. graminearum*)<sup>30</sup> : la fusarénone (ou nivalénol monoacétate), la 7-hydroxyisotrichodermine, la 8-hydroxyisotrichodermine, l'apotrichodiol, la 7,8-dihydroxycalonectrine, la 7,8-dihydroxycalonectrine tétraol, la 15-déacetyl-7,8-dihydroxycalonectrine, le déépoxy-diacétoxyscirpénol, la 5a-hydroxyculmorine et la 14-hydroxyculmorine.

La zéaralénone est une mycotoxine également produite par *F. graminearum*. Les effets œstrogénomimétiques de ce composé induisent chez les animaux des œstrus prolongés ou au contraire absents, une infertilité et une lactation anormale<sup>31</sup>. Chez les mâles, la zéaralénone est responsable d'une diminution du taux de testostérone sérique, du poids des organes reproducteurs et de la spermatogénèse. Une féminisation et une suppression de la libido ont également été observées. AntiBase référence également des dérivés de la zéaralénone parmi les métabolites secondaires produits par *F. graminearum* : la 7'-déhydrozéaralénone, la 5-formylzéaralénone, 8'-épihydroxy-zéaralénone, 8'-hydroxy-zéaralénone, la 4-acétyl cis-zéaralénone et la 4-acétyl zéaralénone.

D'autres métabolites secondaires non réglementés sont également référencés chez *F. graminearum*. La fusarine C est une molécule dont l'ingestion a été associée à l'apparition de papillomes et de carcinomes des voies aérodigestives chez les rongeurs, permettant de le classer parmi les composés possiblement carcinogènes pour l'homme<sup>32</sup> (groupe 2B). L'aurofusarine est un métabolite dont les rôles dans les processus de développement du champignon sont encore mal définis. En revanche, sa structure de type naphtoquinone suggère une action toxique via la production d'ions superoxydes. Le buténolide, la culmorine, les fusariélines, le sambucinol et la chrysogine sont d'autres mycotoxines moins documentées produites par *F. graminearum*. Dans la base de données AntiBase, d'autres métabolites secondaires sont référencés comme étant produits par *F. graminearum* ou *Gibberella zeae* (la forme sexuée de *F. graminearum*)<sup>30</sup> : l'orcinol, la ferricrocine, la triacétylfusarinine, la fusarochromanone, l'anguidine, la monoliformine, et l'acide 4-acétamido-2-butenoïque. La fusaristatine A est également un métabolite secondaire ayant été détecté chez *F. graminearum*<sup>33</sup> et qui a démontré des propriétés inhibitrices de la croissance cellulaire<sup>34</sup>.



Le séquençage de *F. graminearum* suggère qu'une part importante de son métabolome secondaire est aujourd'hui non caractérisée. En effet, sur les 15 PKS identifiées dans son génome, seules six ont été caractérisées du point de vue de leur fonction, et seules trois des 19 NRPS ont été associées à la production de métabolites connus<sup>35,36</sup>.

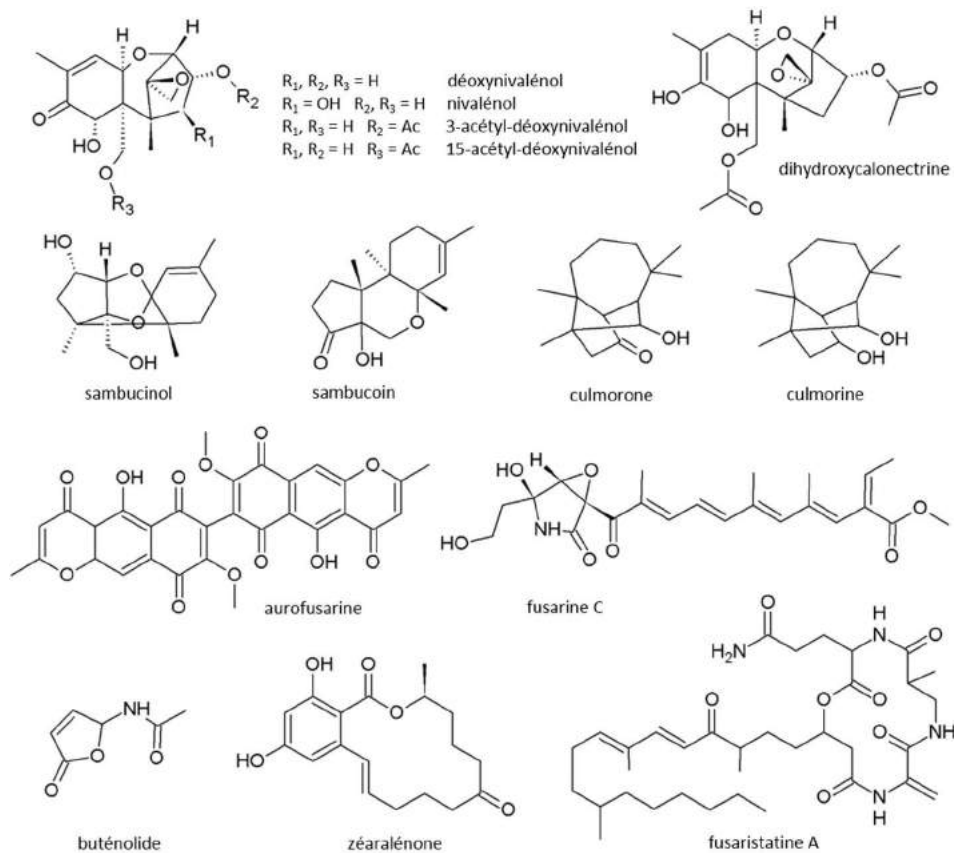


Figure 9. Principaux métabolites secondaires de *F. graminearum*.

## 5. LES FUNGISPORINES

Lors de l'analyse des métabolomes fongiques de *P. verrucosum* et de *P. nordicum*, une famille de métabolites appelés fungisporines a été détectée et a fait l'objet d'une attention particulière (voir Travail expérimental). Cette section est dédiée à la présentation de ces métabolites secondaires.

### 5.1. Une découverte ancienne, une caractérisation récente

La caractérisation du métabolome secondaire de *P. verrucosum* et son analyse par les réseaux moléculaires a permis de mettre en évidence la production de fungisporines par ce champignon filamenteux. Les fungisporines sont des tétrapeptides cycliques observés pour la première fois en 1952 chez neuf espèces de *Penicillium* (*P. notatum*, *P. chrysogenum*, *P. citrinum*, *P. claviforme*, *P. purpurogenum*, *P. herquei*, *P. frequentans*, *P. roseo-purpureum* et *P. brevicaulis*) ainsi que deux espèces d'*Aspergillus* (*A. niger* et *A. versicolor*). La fungisporine A, détectée alors, est considérée comme un octopeptide constitué de résidus de L-phénylalanine et L-valine en quantités équimolaires<sup>37,38</sup>. En 1969, Studer *et al.* pointent du doigt la dimérisation de ce composé et caractérisent précisément la fungisporine A comme le tétrapeptide cyclique *cyclo*(D-Val-L-Val-D-Phe-L-Phe)<sup>39</sup>. L'analyse de la littérature témoigne de l'absence d'étude de ces composés au cours des 40 années qui suivent.

En 2014, lors de l'analyse génomique et transcriptomique de *P. chrysogenum*, Ali *et al.* mettent en évidence la présence d'un gène de NRPS (nommé *hcpA*) fortement exprimé et spécialisé dans la synthèse de tétrapeptides<sup>40</sup> (Tableau 2).

Nom du métabolite	Séquence peptidique	Formule Moléculaire	m/z de [M+H] <sup>+</sup>	Ali <i>et al.</i> 2014 <sup>40</sup>	Klitgaard <i>et al.</i> 2015 <sup>41</sup>	Nielsen <i>et al.</i> 2017 <sup>5</sup>
<b>TETRAPEPTIDES CYCLIQUES</b>						
Fungisporine A	<i>cyclo</i> (FFVV)	C <sub>28</sub> H <sub>36</sub> N <sub>4</sub> O <sub>4</sub>	493.2806	X	X	X
Fungisporine B	<i>cyclo</i> (YFVV)	C <sub>28</sub> H <sub>36</sub> N <sub>4</sub> O <sub>5</sub>	509.2757	X	X	
Fungisporine C	<i>cyclo</i> (YYVV)	C <sub>28</sub> H <sub>36</sub> N <sub>4</sub> O <sub>6</sub>	525.2708		X	
Fungisporine D	<i>cyclo</i> (FWVV)	C <sub>30</sub> H <sub>37</sub> N <sub>5</sub> O <sub>4</sub>	532.2916	X	X	
	<i>cyclo</i> (YWVV)	C <sub>30</sub> H <sub>37</sub> N <sub>5</sub> O <sub>5</sub>	548.2868	X		
	<i>cyclo</i> (FFVI)	C <sub>29</sub> H <sub>38</sub> N <sub>4</sub> O <sub>4</sub>	507.2962	X		
	<i>cyclo</i> (FFIV)	C <sub>29</sub> H <sub>38</sub> N <sub>4</sub> O <sub>4</sub>	507.2962	X		X
	<i>cyclo</i> (YWVI)	C <sub>31</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	562.3017	X		
	<i>cyclo</i> (YWIV)	C <sub>31</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	562.3017	X		
	<i>cyclo</i> (YFVI)	C <sub>29</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	523.2914	X		
	<i>cyclo</i> (YFIV)	C <sub>29</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	523.2914	X		X
<b>TETRAPEPTIDES LINEAIRES</b>						
	FVVF	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	511.2912	X		
	VFFV	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	511.2912	X		X
	FFVV	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	511.2912	X		
	YFVV	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>6</sub>	527.2864	X		
	VYFV	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>6</sub>	527.2864	X		X
	FVVY	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>6</sub>	527.2864	X		
	YWVV	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	566.2972	X		
	VYWV	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	566.2972	X		X
	WVYV	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	566.2972	X		
	VFWV	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	550.3023	X		
	FWVV	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	550.3023	X		X
	WVVF	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	550.3023	X		
	FVIF	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>5</sub>	525.3077	X		
	FIVF	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>5</sub>	525.3077	X		
	FVIY	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>6</sub>	541.3023	X		
	IYFV	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>6</sub>	541.3023	X		
	FIVY	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>6</sub>	541.3023	X		
	VYFI	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>6</sub>	541.3023	X		
<b>METABOLITES ASSOCIES NON IDENTIFIES</b>						
	Indéterminé	C <sub>33</sub> H <sub>44</sub> N <sub>4</sub> O <sub>5</sub>	577.3384		X	
	Indéterminé (prénylé)	C <sub>35</sub> H <sub>39</sub> N <sub>7</sub> O <sub>6</sub>	654.3035		X	
	Indéterminé	C <sub>30</sub> H <sub>31</sub> N <sub>7</sub> O <sub>6</sub>	586.2401		X	
	Indéterminé	C <sub>30</sub> H <sub>37</sub> N <sub>5</sub> O <sub>6</sub>	563.2738			X

Tableau 2. Tétrapeptides cycliques, tétrapeptides linéaires non modifiés et autres composés associés référencés dans la littérature.

La comparaison des profils métaboliques du champignon sauvage et du mutant  $\Delta hcpA$  met en relation ce gène avec la production de 10 tétrapeptides cycliques. Ces composés présentent des structures similaires composées de deux acides aminés aromatiques (phénylalanine, tyrosine et/ou tryptophane) et de deux acides aminés aliphatiques (valine et/ou isoleucine). Dans cette étude, les auteurs soulignent également la présence du gène *hcpA* et de la production de ces 10 métabolites chez *A. niger*. L'année suivante, alors qu'ils étudient chez *A. nidulans* la synthèse de la nidulanine A, Klitgaard *et al.* observent à leur tour la production de fungisporines<sup>41</sup>. En plus de certains métabolites mis en évidence par Ali *et al.*, la structure de la fungisporine C a été caractérisée. L'utilisation des réseaux moléculaires a également permis aux auteurs de souligner la présence de composés inconnus appartenant potentiellement à cette famille de molécules naturelles (Tableau 2). Enfin, lors de l'étude du potentiel métabolique de 24 espèces de *Penicillium*, Nielsen *et al.* ont souligné que le cluster codant pour des NRPS le plus abondant était à l'origine de la synthèse des fungisporines<sup>5</sup>. En effet, la fungisporine A a été détectée chez les 24 espèces étudiées. De par leurs compositions chimiques, sept autres métabolites ont également été annotés en tant que fungisporines chez *P. italicum*, *P. griseofulvum*, *P. expansum*, *P. digitatum*, *P. rubens*, *P. vulpinum* et *P. flavigenum* (Tableau 2).

D'après les auteurs, ces résultats suggèrent que les fungisporines sont des métabolites secondaires produits par la majorité des espèces de *Penicillium*. Dix-huit tétrapeptides linéaires ont également été caractérisés par Ali *et al.* et détectés à travers ces différentes études (Tableau 2)<sup>40</sup>. Comme leurs analogues cycliques, ces métabolites issus du gène *hcpA* sont composés de deux acides aminés aromatiques consécutifs et de deux acides aminés aliphatiques consécutifs.

Lors d'une analyse cinétique de la production des tétrapeptides, il a été observé que les courbes de concentration en peptides cycliques et en peptides linéaires étaient inverses. Ces données suggèrent que les peptides linéaires sont issus de la dégradation des tétrapeptides cycliques par hydrolyse de leurs liaisons peptidiques. La caractérisation des séquences des différents isomères linéaires a démontré que cette hydrolyse n'était possible qu'entre deux acides aminés aromatiques, entre deux acides aminés aliphatiques, ou entre un acide aminé aliphatique et un acide aminé aromatique de son côté N terminal. En effet, aucun peptide comportant un résidu aromatique à son extrémité C terminale et un résidu aliphatique à son extrémité N terminale n'a été observé dans ces études (Figure 10). De plus, considérant la stabilité des tétrapeptides cycliques face aux agents chimiques ou à un stress thermique, Ali *et al.* ont suggéré que ces métabolites linéaires soient issus d'une dégradation enzymatique de leur analogues cycliques<sup>40</sup>. Le fait que ces peptides linéaires peuvent être produits dans le milieu de culture du champignon met en lumière l'intérêt de ces métabolites dans le cadre de l'étude des métabolomes secondaires fongiques. En effet, si ceux-ci avaient été générés lors de leur analyse, ils ne pourraient pas être présentés comme des métabolites secondaires fongiques et seuls leurs peptides cycliques d'origine auraient un intérêt du point de vue fonctionnel. Dans le cadre de cette thèse, les peptides linéaires détectés ont donc été considérés comme des métabolites secondaires à part entière.

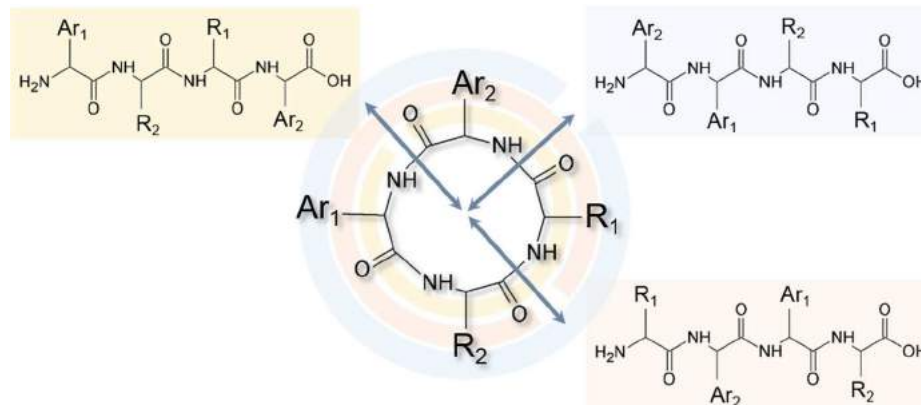


Figure 10. Lien hypothétique entre les tétrapeptides linéaires et les tétrapeptides fongiques observés chez Ali *et al.*<sup>40</sup> Les flèches représentent des dégradations enzymatiques à l'origine de la génération de peptides linéaires. R = groupement aliphatique, Ar = groupement aromatique.

### 5.2. Origines génétiques des fungisporines

L'étude du gène de la NRPS responsable chez *P. chrysogenum* de la synthèse des fungisporines (*hcpA*) par Ali *et al.* a mis en évidence quatre domaines responsables du recrutement des différents substrats, appelés domaines d'adénylation (notés A) (Figure 11)<sup>40</sup>. L'analyse bioinformatique de la structure de ces domaines a supposé l'implication de A1 et A2 dans la prise en charge de phénylalanines et de A3 et A4 dans le recrutement de valines. Néanmoins, les peptides caractérisés comportent également d'autres types d'acides aminés. Est alors suggéré la présence de

microhétérogénéités distinctes entre les modules. Ainsi, le module A1 présente une relative liberté de prise en charge du substrat permettant l'incorporation de phénylalanine mais également de tyrosine, tandis que le domaine A2 permet l'incorporation de phénylalanine et à moindre mesure de tryptophane. Les domaines A3 et A4 sont quant à eux responsables de la prise en charge de valine ainsi que d'autres acides aminés aliphatiques à courte chaîne latérale. Les domaines E présents dans les modules 1 et 3 sont responsables de l'épimérisation des acides aminés activés, expliquant la conformation D des acides aminés incorporés à ces positions. Les domaines PCP et C garantissent quant à eux l'élongation de la séquence (fixation de la chaîne et catalyse de la formation de la liaison peptidique) ainsi que la cyclisation de la molécule. Comme l'organisation linéaire des domaines ne reflète pas l'assemblage linéaire des acides aminés dans le produit final (avec notamment un module 4 entrecoupé de séquences des modules 2 et 3), les auteurs ont déduit que des processus non-linéaires de *trans*-aminoacylation se réalisent probablement grâce à la structure tertiaire de l'enzyme qui permettrait des interactions entre des domaines non consécutifs. Lors de l'étude du métabolome d'*A. nidulans* par Klitgaard *et al.*, il a été démontré que les nidulanines et les fungisporines étaient produites par une même NRPS dont le gène (*nlsA*) présente près de 50% d'homologie avec le gène *hcpA* d'*A. niger*<sup>41</sup>.

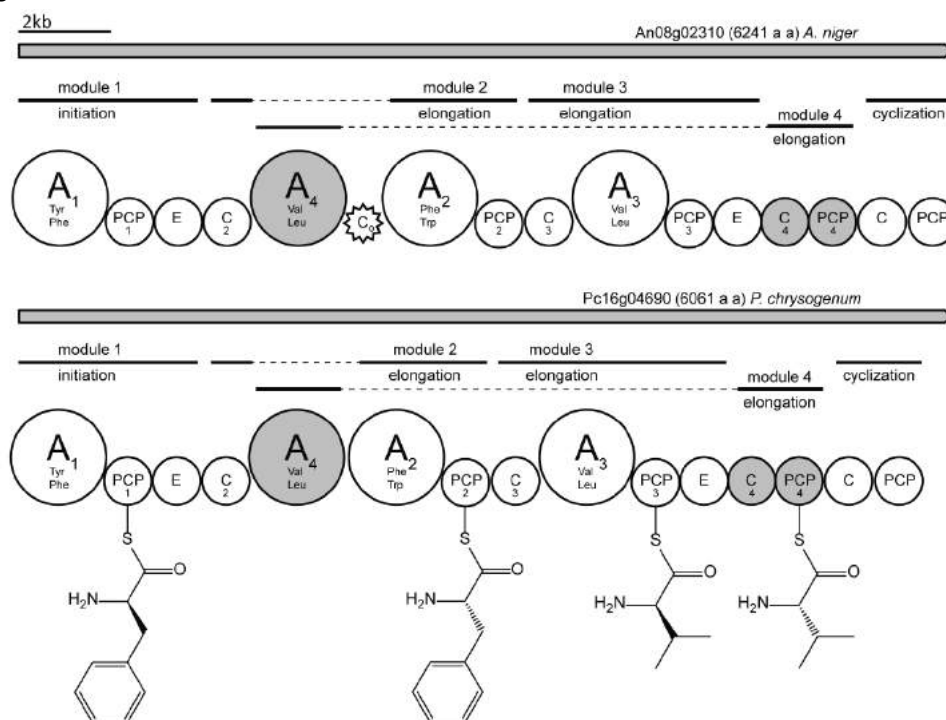


Figure 11. Schématisation des gènes *hcpA* d'*A. niger* (haut) et de *P. chrysogenum* (bas) responsables notamment de la synthèse de fungisporines chez ces champignons. Extrait de Ali *et al.* 2014<sup>40</sup>

### 5.3. Des métabolites d'intérêt

Les fungisporines sont des métabolites secondaires suspectés d'intervenir dans les processus de développement des champignons filamenteux qui les produisent par stimulation de la croissance aérienne du mycélium. En effet, Ali *et al.* ont démontré que la mutation du gène *hcpA* chez *P. chrysogenum* entraîne une modification de l'aspect des cultures et une absence de mycélium aérien (Figure 12)<sup>40</sup>. Ces premiers résultats suggèrent que les fungisporines puissent contaminer les substrats et être liées aux problèmes de destruction des denrées alimentaires par des champignons pathogènes. Suite à ces observations, Ali *et al.* ont suggéré que les fungisporines puissent avoir des fonctions

similaires à celles des hydrophobines<sup>40</sup>. Ces dernières sont impliquées dans les différentes étapes de développement du champignon en permettant l'adhérence du mycélium au substrat, la croissance aérienne et la dissémination des spores. En particulier, il a été démontré que les hydrophobines sont impliquées dans des interactions plante-champignon et que certaines d'entre elles sont responsables du pouvoir pathogène de certains champignons phytopathogènes<sup>42</sup>.

L'intérêt pour les fungisporines étant récent, l'implication de ces métabolites secondaires dans le développement du champignon ainsi que leurs potentiels effets toxiques n'ont pas été d'avantage investigués. Néanmoins, de nombreux tétrapeptides cycliques ou linéaires, composés d'acides aminés standards ou faiblement modifiés, ont démontré des activités toxiques ou des activités à potentielles applications industrielles dans la littérature.

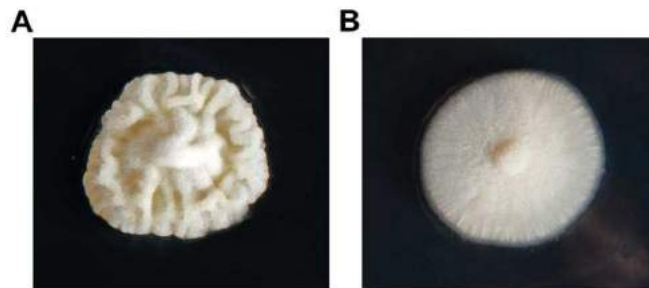


Figure 12. Culture de *P. chrysogenum*. A. Souche sauvage à surface ridée. B. Phénotype lisse du mutant  $\Delta hcpA$ . Source: Ali et al. 2014

En terme de toxicité, des tétrapeptides cycliques fongiques, tels que la dihydrotentoxine (*cyclo*(Leu-NMePhe-Gly-NMeAla)), ont été associés à une phytotoxicité. Bien que de nature biosynthétique différente, les ustiloxines sont des tétrapeptides cycliques modifiés (méthylation, hydroxylation et ajout de norvaline sur la tyrosine) impliqués dans la maladie du faux charbon infectant plus d'un tiers des champs de riz mondiaux<sup>43,44</sup>. Parmi ces métabolites, les ustiloxines A, B et G ont démontré des activités inhibitrices de l'allongement de la radicule et du germe des graines de riz comparables à celle du glyphosate, utilisé comme désherbant<sup>45</sup>. D'autres types de toxicités sont associés à des tétrapeptides cycliques fongiques, comme notamment des activités inhibitrices de tyrosinase par le *cyclo*(PYPV)<sup>46</sup>.

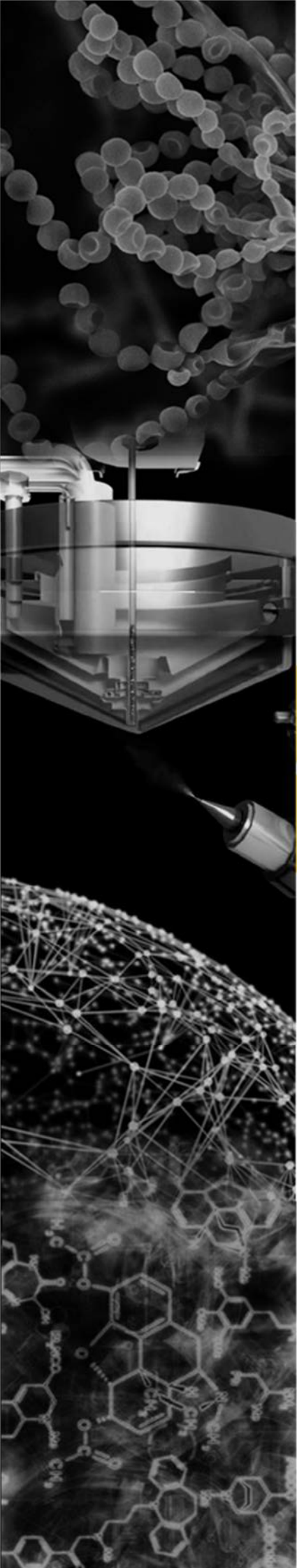
Les tétrapeptides linéaires ou cycliques sont des composés auxquels sont également associés des propriétés pouvant être utilisées à des fins médicales ou industrielles. En effet, le tétrapeptide linéaire D-Phe-L-Val-D-Val-L-Tyr (potentiellement issu de l'hydrolyse de la fungisporine B) caractérisé chez *P. canescens* a démontré une activité antifongique comparable à celle du produit phytosanitaire Benomyl<sup>47</sup>. De plus, les tétrapeptides cycliques *cyclo*(N-MePhe-Ile)<sub>2</sub>, *cyclo*(N-MePhe-Val)<sub>2</sub>, et *cyclo*(N-MePhe-Val-N-MePhe-Ile) découverts chez *Onychochola sclerotica* ont été associés à des propriétés bloquantes vis-à-vis des canaux cardiaques, faisant de ces molécules un point de départ pour la recherche de nouveaux traitements contre l'arythmie, l'angor ou l'hypertension. Inhibant spécifiquement les canaux sodiques, ces métabolites présenteraient une toxicité inférieure aux bloqueurs de canaux potassiques disponibles sur le marché<sup>48</sup>. La forte similarité entre les fungisporines et ces composés cycliques suggère que les fungisporines puissent présenter des propriétés similaires<sup>40</sup>.

## 6. RÉFÉRENCES DE L'INTRODUCTION

---

1. Bennett JW, Klich M, Mycotoxins M. Mycotoxins. *Clin Microbiol Rev.* 2003;16(3):497-516.
2. Fisher MC, Henk DA, Briggs CJ, et al. Emerging fungal threats to animal, plant and ecosystem health. *Nature.* 2013;484(7393):1-18.
3. Meyer V, Andersen MR, Brakhage AA, et al. Current challenges of research on filamentous fungi in relation to human welfare and a sustainable bio-economy: a white paper. *Fungal Biol Biotechnol.* 2016;3(1):6.
4. Bérdy J. Thoughts and facts about antibiotics: Where we are now and where we are heading. *J Antibiot (Tokyo).* 2012;65(8):441-441.
5. Nielsen JC, Grijsseels S, Prigent S, et al. Global analysis of biosynthetic gene clusters reveals vast potential of secondary metabolite production in *Penicillium* species. *Nat Microbiol.* 2017;2(April):17044.
6. Cano PM, Jamin EL, Tadrist S, et al. New untargeted metabolic profiling combining mass spectrometry and isotopic labeling: Application on *Aspergillus fumigatus* grown on wheat. *Anal Chem.* 2013;85(17):8412-8420.
7. Samson RA, Yilmaz N, Houbraken J, et al. Phylogeny and nomenclature of the genus *Talaromyces* and taxa accommodated in *Penicillium* subgenus *Biverticillium*. *Stud Mycol.* 2011;70:159-183.
8. Gouvernement-du-Québec. INSPQ. *Penicillium* spp. July 2016. <https://www.inspq.qc.ca/moisissures/fiches/penicillium-spp>. Accessed June 1, 2017.
9. Université-de-Brest. *Penicillium*.
10. Frisvad JC, Smedsgaard J, Larsen TO, Samson RA. Mycotoxins, drugs and other extrolites produced by species in *Penicillium* subgenus *Penicillium*. *Stud Mycol.* 2004;49:201-241.
11. Frisvad JC, Samson RA. Polyphasic taxonomy of *Penicillium* subgenus *Penicillium*: A guide to identification of food and air-borne terverticillate *Penicillia* and their mycotoxins. *Stud Mycol.* 2004;49:1-174.
12. JECFA JF. *WHO Technical Report Series.*; 2007.
13. Elmholt S, Rasmussen PH. *Penicillium verrucosum* occurrence and ochratoxin A contents in organically cultivated grain with special reference to ancient wheat types and drying practice. *Mycopathologia.* 2005;159(3):421-432.
14. Ciegler A, Fennell DI, Sansing GA, Detroy RW, Bennett GA. Mycotoxin-producing strains of *Penicillium viridicatum*: Classification into subgroups. *Appl Environ Microbiol.* 1973;26(3):271-278.
15. Pitt JI. *Penicillium viridicatum*, *Penicillium verrucosum*, and production of ochratoxin A. *Appl Environ Microbiol.* 1987;53(2):266-269.
16. Larsen TO, Svendsen A, Smedsgaard J. Biochemical characterization of ochratoxin A-producing strains of the genus *Penicillium*. *Appl Environ Microbiol.* 2001;67(8):3630-3635.
17. Samson RA, Seifert KA, Kuijpers AF., Houbraken JAM., Frisvad JC. Phylogenetic analysis of *Penicillium* subgenus *Penicillium* using partial  $\beta$ -tubulin sequence. *Stud Mycol.* 2004;49:175-200.
18. Cabanes FJ, Bragulat MR, Castella G. Ochratoxin A producing species in the genus *Penicillium*. *Toxins (Basel).* 2010;2(5):1111-1120.
19. Frisvad JC, Filtenborg O. Classification of terverticillate *penicillia* based on profiles of mycotoxins and other secondary metabolites. *Appl Environ Microbiol.* 1983;46(6):1301-1310.
20. Rahbaek L, Sperry S, Frisvad JC, Larsen TO. PC-2, LL-P888gamma and some novel analogue alpha-pyrone from *Penicillium nordicum*, *P. verrucosum* and *P. olsonii*. *Biochem Syst Ecol.* 2003;31(3):313-317.
21. el Khoury AE, Atoui A. Ochratoxin a: General overview and actual molecular status. *Toxins (Basel).* 2010;2(4):461-493.
22. International Agency for Research on Cancer (IARC). Ochratoxin A. *IARC Monogr Eval Carcinog Risk Chem to Humans.* 1993;56:489-521.
23. Frisvad JC, Lund F, Elmholt S. Ochratoxin A producing *Penicillium verrucosum* isolates from cereals reveal large AFLP fingerprinting variability. *J Appl Microbiol.* 2005;98(3):684-692. doi:72.2004.02509.x.
24. Flajs D, Peraica M. Toxicological properties of citrinin. *Arh Hig Rada Toksikol.* 2009;60:457-464.
25. Leong SL, Schnürer J, Broberg A. Verrucine F, a quinazoline from *Penicillium verrucosum*. *J Nat Prod.* 2008;71(8):1455-1457.
26. Malik RK, Engel G, Teuber M. Applied Microbiology Biotechnology Effect of some nutrients on the production of cyclopiazonic acid by *Penicillium verrucosum* var. *cyclopium*. *Appl Microbiol Biotechnol.* 1986;24:71-74.

27. Leslie J, Summerell B. *The Fusarium Laboratory Manual*. Blackwell. Oxford, UK; 2006.
28. Pinton P, Oswald IP. Trichothecenes on the Intestine : A Review. *Toxins (Basel)*. 2014;6:1615-1643.
29. Rocha O, Ansari K, Doohan FM. Effects of trichothecene mycotoxins on eukaryotic cells: A review. *Food Addit Contam*. 2005;22(4):369-378.
30. Laatsch. Laatsch 2012. <http://www.wiley-vch.de/stmdata/antibase.php>).
31. Zinedine A, Soriano JM, Moltó JC, Mañes J. Review on the toxicity, occurrence, metabolism, detoxification, regulations and intake of zearalenone: An oestrogenic mycotoxin. *Food Chem Toxicol*. 2007;45(1):1-18.
32. (IARC) IA for R on C. *Report*. Vol 56.; 1997.
33. Sørensen JL, Sondergaard TE, Covarelli L, et al. Identification of the biosynthetic gene clusters for the lipopeptides fusaristatin A and W493 B in *Fusarium graminearum* and *F. pseudograminearum*. *J Nat Prod*. 2014;77(12):2619-2625.
34. Shiono Y, Tsuchinari M, Shimanuki K, et al. Fusaristatins A and B, two new cyclic lipopeptides from an endophytic *Fusarium* sp. *J Antibiot (Tokyo)*. 2007;60(5):309-316.
35. Gaffoor I, Brown DW, Plattner R, Proctor RH, Qi W, Trail F. Functional analysis of the polyketide synthase genes in the filamentous fungus *Gibberella zeae* (anamorph *Fusarium graminearum*). *Eukaryot Cell*. 2005;4(11):1926-1933.
36. Hansen F, Sorensen J, Giese H, Sondergaard T, Frandsen R. Quick guide to polyketide synthase and nonribosomal synthetase genes in *Fusarium*. *Int J Food Microbiol*. 2012;155:128-136.
37. Miyao K. 14. Studies on Fungisporin. Part 2. *J Agric Chem Soc Japan*. 1955;19(1):86-91.
38. Miyao K. The Structure of Fungisporin. *Bull Agric Chem Soc Japan*. 1960;24(1):23-30.
39. Studer RO. Synthesis and structure of Fungisporin. *Experientia*. 1969;25(9):899.
40. Ali H, Ries MI, Lankhorst PP, et al. A non-canonical NRPS is involved in the synthesis of fungisporin and related hydrophobic cyclic tetrapeptides in *Penicillium chrysogenum*. *PLoS One*. 2014;9(6):e98212.
41. Klitgaard A, Nielsen JB, Frandsen RJN, Andersen MR, Nielsen KF. Combining stable isotope labeling and molecular networking for biosynthetic pathway characterization. *Anal Chem*. 2015;87(13):6520-6526.
42. Whiteford JR, Spanu PD. Hydrophobins and the interactions between fungi and plants. *Mol Plant Pathol*. 2002;3(5):391-400.
43. Umemura M, Nagano N, Koike H, et al. Characterization of the biosynthetic gene cluster for the ribosomally synthesized cyclic peptide ustiloxin B in *Aspergillus flavus*. *Fungal Genet Biol*. 2014;68:23-30.
44. Koiso Y, Natori M, Iwasaki S, et al. Ustiloxin: a phytotoxin and a mycotoxin from false smut balls on rice panicles. *Tetrahedron Lett*. 1992;33(29):4157-4160.
45. Wang X, Wang J, Lai D, et al. Ustiloxin G, a New Cyclopeptide Mycotoxin from Rice False Smut Balls. *Toxins (Basel)*. 2017;9(2):54.
46. Kawagishi. A novel cyclotetrapeptide produced by *Lactobacillus helveticus* as a tyrosinase inhibitor. *Tetrahedron Lett*. 1993;34(21):3439-3440.
47. Bertinetti B V., Peña NI, Cabrera GM. An antifungal tetrapeptide from the culture of *Penicillium canescens*. *Chem Biodivers*. 2009;6(8):1178-1184.
48. Pérez-Victoria I, Martín J, González-Menéndez V, et al. Isolation and structural elucidation of cyclic tetrapeptides from *Onychocola sclerotica*. *J Nat Prod*. 2012;75(6):1210-1214.



# Matériels & Méthodes

« On ne peut se passer d'une méthode pour se mettre en quête de la vérité des choses. »

*René Descartes*





Ce chapitre présente la méthode de caractérisation des métabolites secondaires utilisée pour l'étude de *Penicillium* et *Fusarium*. Dans un premier temps, la génération des données est détaillée. Puis, les outils utilisés pour l'interprétation des résultats sont présentés. Dans le cadre de cette thèse, des fungisporines et certains de leurs dérivés ont été mis en évidence et ont fait l'objet d'une élucidation structurale (voir Travail expérimental). Ces molécules étant des tétrapeptides cycliques et linéaires, une troisième section du chapitre Matériel & méthode est dédiée au séquençage peptidique *de novo*.

## 1. LA GÉNÉRATION DES DONNÉES

### 1.1. La méthode analytique

#### 1.1.1. Le principe

La stratégie analytique employée dans le cadre de cette thèse a été précédemment développée au sein de l'unité Toxalim de l'INRA de Toulouse, en partenariat avec le Groupe de Recherches Appliquées en Phytotechnologies du Commissariat à l'Énergie Atomique et aux énergies alternatives (CEA) de Cadarache. Elle repose sur un double marquage du métabolome secondaire du champignon étudié et sur son analyse par des méthodes de métabolomique non ciblée (Figure 13). Cette méthode a été validée par l'analyse du modèle *A. fumigatus* dont le métabolome secondaire est l'un des mieux référencés<sup>1</sup>.

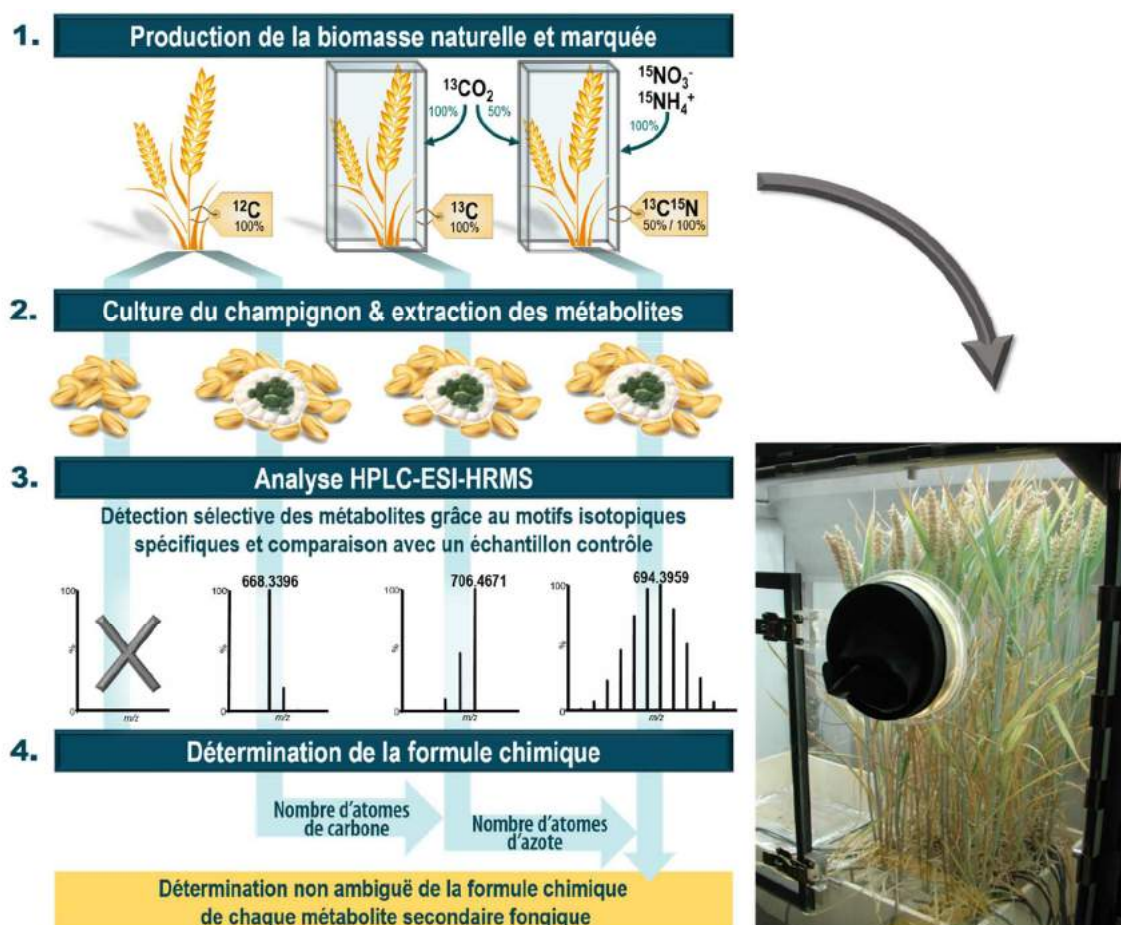


Figure 13. Stratégie analytique mise en œuvre pour l'analyse des métabolomes secondaires fongiques par double marquage isotopique et spectrométrie de masse haute résolution.

### 1.1.2. La métabolomique non ciblée

La métabolomique est une discipline dédiée à l'analyse de l'ensemble des métabolites présents dans un milieu biologique (le métabolome) et dont la masse moléculaire est généralement inférieure à 1 500 Da. L'analyse des métabolites peut être effectuée de manière ciblée ou non ciblée. Dans le premier cas, les analyses sont dirigées spécifiquement vers un certain nombre de molécules connues, en vue de leur dosage par exemple. Par opposition, la métabolomique non ciblée est l'étude de l'ensemble des composés d'un échantillon biologique sans *a priori*. Pour la recherche de nouveaux métabolites secondaires fongiques, l'une des techniques de métabolomique non ciblée consiste en l'analyse de l'empreinte métabolique d'un extrait, obtenue par spectrométrie de masse couplée à la chromatographie. Le spectre de masse résultant est alors composé de signaux représentant toutes les molécules ionisées dans les conditions de l'analyse. Selon les techniques utilisées (solvant d'extraction, type de chromatographie, type d'ionisation...), le type de molécules détectées pourra être différent. Dans le cadre de cette thèse, une méthode couplant la chromatographie liquide haute performance (HPLC) et la spectrométrie de masse haute résolution (HRMS) avec une ionisation de type électrospray a été utilisée. En raison de la complexité des empreintes obtenues, l'analyse des données métabolomiques est généralement assistée par des logiciels tels que XCMS<sup>®</sup> et MZmine<sup>®</sup>. Ces derniers réalisent notamment l'étape d'extraction des pics qui consiste à sélectionner les signaux issus de molécules et à soustraire le bruit de fond. Néanmoins, de telles analyses ne permettent pas de différencier les métabolites des contaminants générés lors du traitement des échantillons. Afin de surpasser cette limite, la méthode utilisée dans le cadre de cette thèse s'appuie sur l'utilisation d'un marquage isotopique.

### 1.1.3. Marquage isotopique du métabolome secondaire fongique

L'unique source de carbone et d'azote d'un champignon provient de son substrat. La mise en contact d'une moisissure avec des atomes de carbone et d'azote marqués par des isotopes stables permet alors de marquer le champignon et donc ses métabolites secondaires. Selon la spécificité de l'unité marquée, la gamme de métabolites marqués est plus ou moins large. Ainsi, l'incubation de moisissures avec des acides aminés marqués ou tout autre précurseur spécifique permet de distinguer certaines familles de composés au sein du métabolome secondaire (Figure 14, gauche). À plus large échelle, le marquage de précurseur non spécifique, source de carbone et d'azote, permet de marquer la totalité des molécules produites par le champignon. Ce type de marquages non spécifiques permet alors de distinguer l'ensemble des molécules fongiques parmi les contaminants formés lors des différentes étapes de l'analyse (Figure 14, droite).

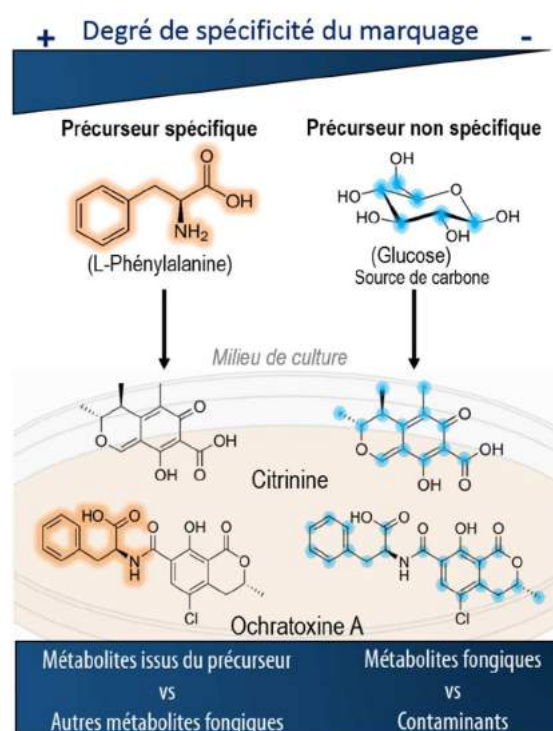


Figure 14. Cibles d'un marquage isotopique selon le degré de spécificité du précurseur marqué.

En vue de l'analyse globale des métabolomes fongiques, la méthode utilisée dans le contexte de cette thèse est fondée sur un marquage non spécifique des métabolites secondaires. Cette stratégie permet de distinguer les molécules du champignon et fournit des informations quant à la composition chimique de chacune d'elles. En effet, la comparaison de la masse d'un métabolite marqué au carbone 13 avec celle de ce même métabolite non marqué permet de connaître la quantité de carbone qu'il contient. Cette information permet de préciser la formule chimique de ce composé parmi les multiples possibilités en accord avec sa masse exacte. Bien que le calcul du nombre d'atomes de carbone puisse être suffisant pour déterminer la formule chimique des petits métabolites secondaires (<400 Da), plusieurs possibilités peuvent subsister, même avec une haute précision de mesure de masse (<5 ppm). Pour cette raison, un second marquage isotopique (à l'azote 15) est intégré dans la méthode développée au sein du laboratoire. Ainsi, en plus du nombre d'atomes de carbone, la connaissance de la composition en azote permet de déterminer de manière non ambiguë la formule chimique de tous les métabolites secondaires fongiques.

Le métabolome secondaire des champignons permet à ces organismes d'adapter leur croissance en fonction de leur environnement. Le principe appelé OSMAC (One Strain MANY Compounds – une souche, de nombreux composés) explique pourquoi tout infime changement du milieu de croissance d'une moisissure peut entraîner une modification importante de son métabolome secondaire. Il est donc crucial d'étudier les moisissures dans leur milieu naturel, en particulier dans le cadre d'analyses visant à comprendre leurs mécanismes d'infestation des denrées alimentaires. Pour ces raisons, la méthode utilisée pour l'analyse de *P. verrucosum* et de *F. graminearum* consiste en la croissance de ces champignons sur grains de blé. Trois blés différemment marqués ont donc été générés grâce à leur croissance dans des chambres hermétiques (photo Figure 13) alimentées avec du CO<sub>2</sub> marqué ou non avec des atomes de carbone 13 et avec des nutriments (NO<sub>3</sub><sup>-</sup> et NH<sub>4</sub><sup>+</sup>) marqués ou non avec des atomes d'azote 15 : (i) un blé naturel (contenant 99% d'isotopes <sup>12</sup>C), (ii) un « blé <sup>13</sup>C » (contenant 97% d'isotopes <sup>13</sup>C) et (iii) un « blé <sup>13</sup>C<sup>15</sup>N » (contenant 97% d'isotopes <sup>15</sup>N et 53% d'isotopes <sup>13</sup>C). Les grains de blé récoltés ont ensuite été utilisés comme unique source d'atomes de carbone et d'azote par la moisissure. Pour chaque souche étudiée, trois cultures ont été réalisées en parallèle sur chacun des types de grains. Le marquage par isotopes stables n'affectant pas les processus de biotransformation, trois métabolomes secondaires composés des mêmes métabolites secondaires mais différemment marqués ont été obtenus au terme de ces expérimentations. Un échantillon contrôle constitué de blé non infesté a été disposé dans les mêmes conditions que les trois cultures pour permettre de différencier les métabolites du blé et des métabolites fongiques au sein des composés marqués (Figure 13).

### 1.1.4. Analyse des métabolomes par HPLC-HRMS

Lors d'une analyse métabolomique non ciblée, le type de molécules détectées dépend des techniques d'extraction et de détection employées. Aucune méthode universelle ne permettant l'observation simultanée de l'ensemble des molécules présentes dans un échantillon biologique, le choix de ces techniques est donc orienté en fonction des classes de composés étudiés. Dans cette étude, les paramètres analytiques ont été choisis de manière à pouvoir détecter le maximum de métabolites secondaires d'une part, et à pouvoir comparer les données obtenues avec celles de la littérature d'autre part. Dans cette section, les paramètres d'extraction, de séparation chromatographique et d'analyse par spectrométrie de masse sont justifiés (pour plus de détails, voir Article 2).

Après culture des champignons sur grains de blé, les métabolites secondaires synthétisés par le micro-organisme ont été extraits à l'acétate d'éthyle pendant 72 heures. Ce solvant est peu toxique et présente une polarité moyenne lui permettant d'extraire un large spectre de composés. Sa volatilité importante facilite les processus d'évaporation de solvant qui suivent l'étape d'extraction. Les métabolites secondaires fongiques ainsi extraits de chacune des cultures ont ensuite été repris dans du méthanol. Il est généralement préférable de suspendre les molécules à analyser par chromatographie liquide dans les conditions initiales de solvants. Dans le cadre de cette thèse, des problèmes de solubilisation de certains composés dans un mélange eau/acétonitrile (80%/20%) ou dans l'acétonitrile pur ont été rencontrés, justifiant ainsi l'emploi de méthanol.

L'analyse métabolomique des composés fongiques a été effectuée par HPLC-HRMS. La séparation chromatographique a été réalisée par chromatographie liquide en phase inverse grâce à une colonne C18. Un gradient de 0,1% d'acide formique dans l'eau (solvant A) et de 100% d'acétonitrile (solvant B) a permis une bonne séparation des métabolites secondaires. La détection des composés a été réalisée par spectrométrie de masse haute résolution à l'aide d'un LTQ Orbitrap XL. Puisqu'elle permet une ionisation douce de la plupart des métabolites secondaires, une source électrospray à température ambiante a été choisie. Deux injections, l'une en mode positif (ESI<sup>+</sup>) et l'autre en mode négatif (ESI<sup>-</sup>), ont été réalisées séparément afin de permettre la détection d'un maximum de composés. La détection des ions a été réalisée en haute résolution dans l'Orbitrap (Figure 15).

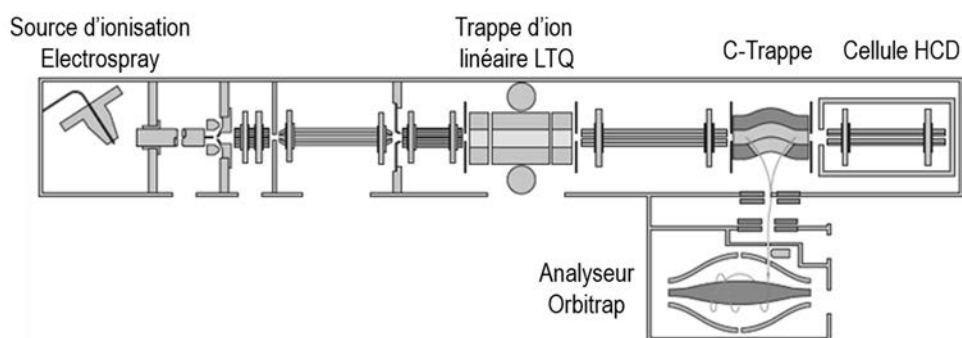


Figure 15. Schéma d'un spectromètre de masse LTQ-Orbitrap. Source : Thermo Scientific.

## 1.2. Analyse des données

### 1.2.1. Extraction des pics

L'analyse de données de spectrométrie de masse en métabolomique non ciblée nécessite une étape d'extraction des pics qui consiste en la détection des signaux dans les différents spectres de masse. Dans le cadre de l'étude présentée ici, cette étape a été facilitée par l'utilisation des deux marquages isotopiques qui permettent une détection spécifique des métabolites secondaires parmi les contaminants et le bruit de fond. L'étape d'extraction des pics a été réalisée en quatre temps (Figure 16).

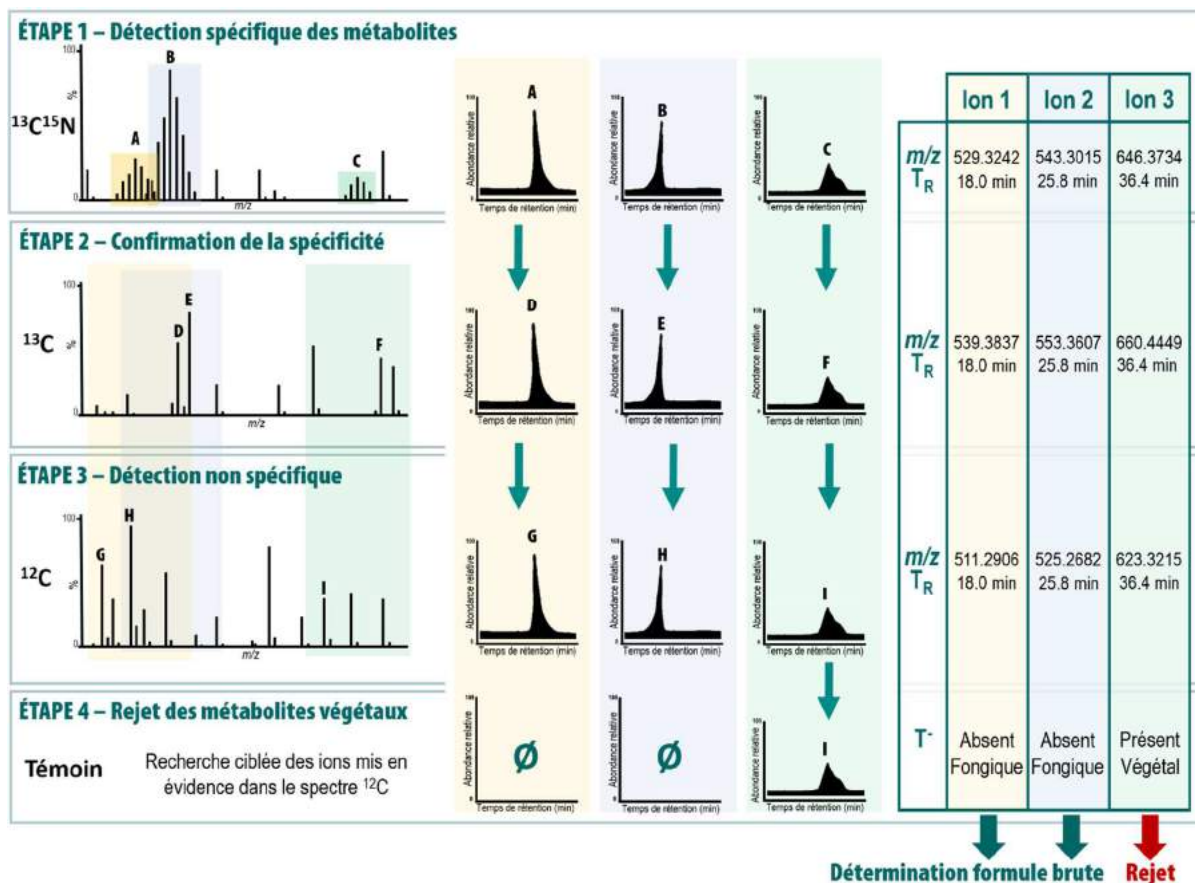


Figure 16. Protocole d'extraction des pics permettant la détection spécifique des métabolites marqués à partir des spectres issus de cultures  $^{13}\text{C}^{15}\text{N}$  et  $^{13}\text{C}$  (étapes 1 et 2), la détermination de l'ion équivalent dans le spectre issu de la culture  $^{12}\text{C}$  (étape 3) et le rejet des molécules d'origine végétale (étape 4).

#### ÉTAPE 1. Détection spécifique des métabolites secondaires

Dans un premier temps, l'analyse des spectres de masse  $^{13}\text{C}^{15}\text{N}$  issus de la culture du champignon sur le blé doublement marqué et obtenus en ionisation positive et négative, a permis une détection spécifique et facilitée des métabolites grâce à leur détection sous forme de massifs isotopiques particuliers (Figure 17). Le motif de type gaussien est généré par le marquage à 53% en atomes de  $^{13}\text{C}$ . Chacun des pics du massif correspond au nombre d'atomes de carbones 13 incorporés dans la molécule au moment de sa synthèse, le pic central possédant environ la moitié de ses atomes de carbones marqués (Figure 18A).

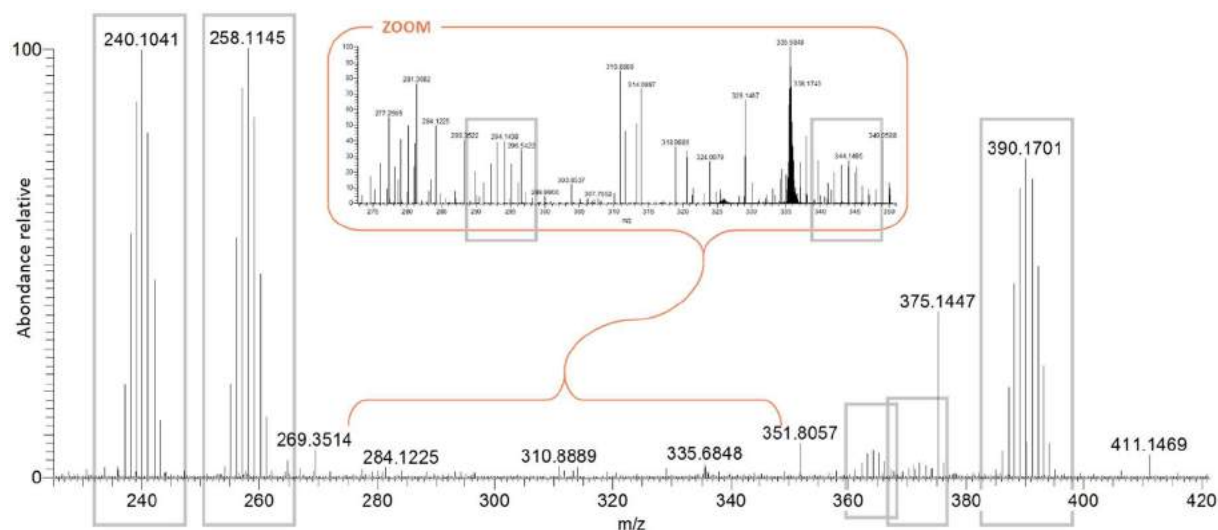


Figure 17. Extrait du spectre de masse issu de l'analyse du métabolome secondaire de *P. verrucosum* cultivé sur blé  $^{13}\text{C}^{15}\text{N}$  (50%  $^{13}\text{C}$ , 100%  $^{15}\text{N}$ ) (spectre moyen de 18 à 22 min,  $m/z$  de 220 à 420). Les ions issus de métabolites secondaires sont spécifiquement et facilement détectés grâce à leurs massifs isotopiques particuliers (encadrés).

### ÉTAPE 2. Confirmation de la spécificité

Dans un second temps, les spectres de masse  $^{13}\text{C}$  issus de la culture du champignon sur blé marqué avec des atomes  $^{13}\text{C}$  et générés en ionisation positive et négative ont été analysés. Cette étape permet de mettre en évidence les métabolites détectés lors de la première étape mais ayant incorporé cette fois-ci des atomes  $^{13}\text{C}$  uniquement, et des atomes d'azote naturels. Le marquage des molécules par des isotopes stables ne modifiant pas les propriétés chromatographiques des métabolites, les composés marqués avec des atomes  $^{13}\text{C}$  présentent les mêmes profils d'éluion chromatographique que leurs équivalents  $^{13}\text{C}^{15}\text{N}$ . Seul leurs rapports  $m/z$  et leurs massifs isotopiques observés en spectrométrie de masse sont modifiés.

Afin d'établir une équivalence entre les métabolites recensés lors de l'étape 1 et les métabolites observés dans les spectres de masse  $^{13}\text{C}$ , les profils d'éluion chromatographique ont été comparés. Les métabolites marqués à 97% par du  $^{13}\text{C}$  ont également été détectés de manière spécifique grâce aux atomes de  $^{12}\text{C}$  résiduels (Figure 18B). Ceci permet d'une part la confirmation de l'origine du métabolite et d'autre part la détection spécifique de composés qui n'auraient pas été détectés lors de la première étape.

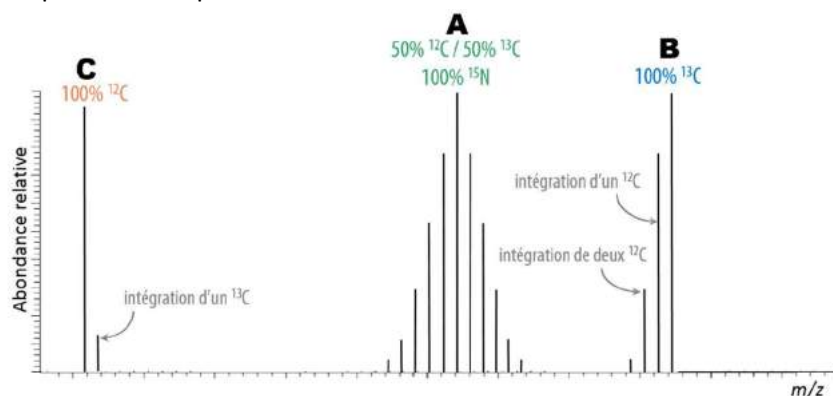


Figure 18. Superposition des massifs isotopiques observés sur les spectres de masse d'un même métabolite marqué 53%  $^{13}\text{C}$  97%  $^{15}\text{N}$  (A), 97%  $^{13}\text{C}$  (B) et 99%  $^{12}\text{C}$  (C). Les massifs isotopiques sont générés selon la probabilité d'intégration des différents isotopes dans la molécule.

### ÉTAPE 3. Détection non spécifique

La troisième étape de l'extraction des pics consiste en la détection des métabolites dans les spectres de masse issus de la culture du champignon sur blé naturel. Les métabolites non marqués ont été recherchés aux temps de rétention observés à partir des analyses des cultures sur les milieux marqués. Les correspondances entre les ions non marqués et les ions marqués ont été confirmées tout d'abord par la comparaison de leurs profils chromatographiques, puis par la comparaison de leurs spectres MS/MS lorsque cela est possible (Figure 19 et voir section 1.3.1). Contrairement aux étapes précédentes, la détection des métabolites non marqués n'est pas spécifique et ne permet pas de distinguer les métabolites secondaires fongiques des contaminants (Figure 18C).

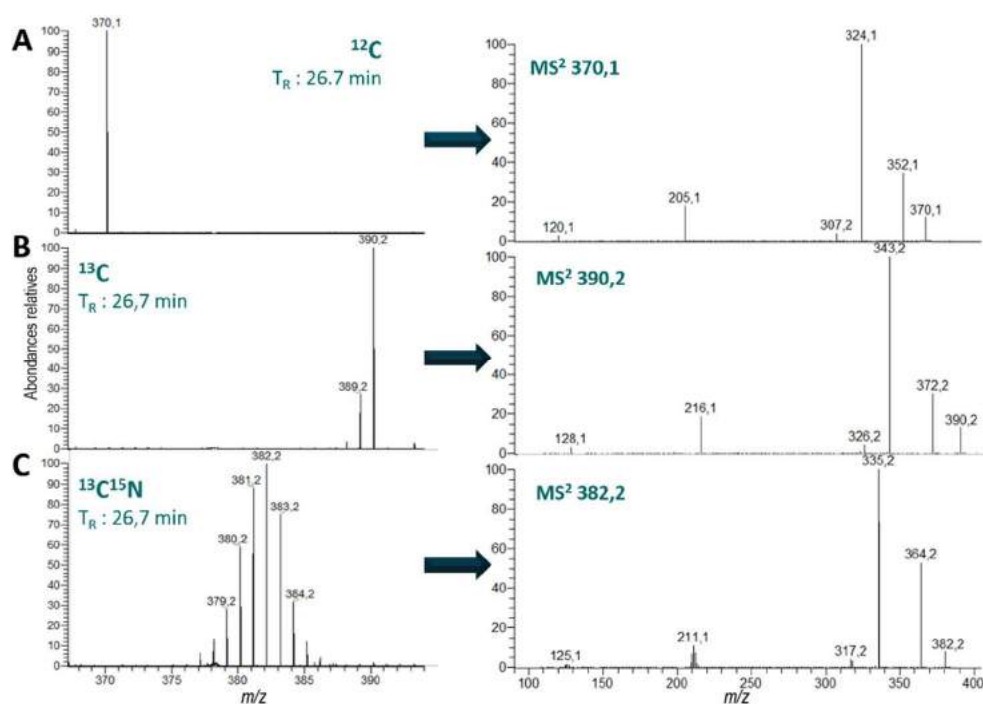


Figure 19. Comparaison des spectres MS/MS obtenus par acquisition en mode data-dépendant des trois ions moléculaires détectés dans les spectres de masse  $^{12}\text{C}$  (A),  $^{13}\text{C}$  (B) et  $^{13}\text{C}^{15}\text{N}$  (C), d'un métabolite équivalent.

### ÉTAPE 4. Rejet des métabolites végétaux

La dernière étape de l'extraction des pics consiste en la vérification de l'origine fongique des métabolites analysés dans cette étude. Pour cela, chacun des métabolites détectés lors des étapes précédentes a été recherché dans le spectre de masse de l'échantillon contrôle (blé non contaminé). Les composés observés dans le témoin ne sont pas issus de l'infestation du blé par le champignon et n'ont donc pas été considérés dans la suite de l'étude.

#### 1.2.2. Suppression de la redondance d'information

Suite à l'extraction des pics, une analyse de la redondance de l'information permet de ne prendre en compte qu'un seul signal par métabolite secondaire afin de ne pas surestimer les capacités métaboliques du champignon étudié. En effet, lors de l'ionisation par électrospray, un métabolite peut être détecté sous plusieurs formes : protonée(s) ; déprotonée(s) adduits avec un(des) cation(s)/anion(s) ou une(des) molécule(s) coéluées (solvant, ammoniac, acide) ; multimères ou fragments produits dans la source d'ionisation. Deux ions présentant des profils chromatographiques



parfaitement identiques ont été considérés comme issus d'un même métabolite secondaire. L'écart de masse entre ces ions a permis de déterminer l'ion à conserver (ion moléculaire de préférence) et les ions redondants à rejeter. Une liste complémentaire d'adduits pouvant être observés est présentée en Tableau 3.

ESI <sup>+</sup>		ESI <sup>-</sup>	
Écart de masse (uma)	Adduit	Écart de masse (uma)	Adduit
<i>m/z<sub>MH+</sub></i>	<b>[M+H]<sup>+</sup></b>	<i>m/z<sub>MH+</sub></i>	<b>[M-H]<sup>-</sup></b>
+ 17,0265	[M+NH <sub>4</sub> ] <sup>+</sup>	- 18,0111	[M-H <sub>2</sub> O-H] <sup>-</sup>
- 18,0153	[M+H-H <sub>2</sub> O] <sup>+</sup>	+ 21,9819	[M+Na-2H] <sup>-</sup>
+ 21,9819	[M+Na] <sup>+</sup>	+ 35,9766	[M+Cl] <sup>-</sup>
+ 32,0262	[M+CH <sub>3</sub> OH+H] <sup>+</sup>	+ 37,9558	[M+K-2H] <sup>-</sup>
+ 37,9558	[M+K] <sup>+</sup>	+ 46,0054	[M+AF-H] <sup>-</sup>
+ 41,0265	[M+ACN+H] <sup>+</sup>	+ 67,9874	[M+AF+Na-2H] <sup>-</sup>
+ 43,9638	[M+2Na-H] <sup>+</sup>	+ 79,9261	[M+Br] <sup>-</sup>
+ 63,0084	[M+ACN+Na] <sup>+</sup>		
+ 75,9117	[M+2K-H] <sup>+</sup>		
+ 82,0530	[M+2ACN+H] <sup>+</sup>		
<b>[2 <i>m/z<sub>MH+</sub></i> - H]<sup>+</sup></b>	<b>[2M+H]<sup>+</sup></b>	<b>[2 <i>m/z<sub>MH+</sub></i> + H]<sup>-</sup></b>	<b>[2M-H]<sup>-</sup></b>
+ 17,0265	[2M+NH <sub>4</sub> ] <sup>+</sup>	+ 46,0054	[2M+AF-H] <sup>-</sup>
+ 21,9819	[2M+Na] <sup>+</sup>		
+ 37,9558	[2M+K] <sup>+</sup>		
+ 41,0265	[2M+ACN+H] <sup>+</sup>		
+ 63,0084	[2M+ACN+Na] <sup>+</sup>		

Tableau 3. Écarts de masse théoriques entre l'ion moléculaire ou le dimère et les adduits correspondants. ACN = acétonitrile ; AF = acide formique

### 1.2.3. Détermination des formules chimiques

Suite à l'étape d'extraction des pics, les formules brutes des métabolites secondaires fongiques ont été déterminées sans ambiguïté. Pour un métabolite donné, les masses mesurées des trois ions correspondants, détectés respectivement dans le spectre de masse non marqué, le spectre de masse <sup>13</sup>C et le spectre de masse <sup>13</sup>C<sup>15</sup>N ont en effet permis de connaître le nombre d'atomes de carbone et d'atomes d'azote que comporte la molécule. La différence de masse entre un atome de carbone 12 et un atome de carbone 13 étant d'une seule unité de masse, le nombre de carbones (N<sub>C</sub>) a été obtenu par la soustraction du *m/z* de l'ion entièrement marqué au <sup>13</sup>C (*m/z*<sub>13C</sub>) et du *m/z* de l'ion non marqué (*m/z*<sub>12C</sub>).

$$N_C = m/z_{13C} - m/z_{12C}$$

La quantité d'azote (N<sub>N</sub>) a quant à elle été déterminée à partir du *m/z* de l'ion doublement marqué (*m/z*<sub>13C15N</sub>) par la formule suivante :

$$N_N = m/z_{13C15N} - m/z_{12C} - \frac{N_C}{2}$$

Lorsque le nombre d'atomes de carbone est impair, le nombre d'azotes correspond à l'un des deux entiers qui encadrent le résultat de cette équation et peut être déterminé grâce à la règle de l'azote. La connaissance de la composition en carbone et azote du métabolite a alors permis de filtrer les différentes possibilités de formules brutes pour chacune des trois masses exactes. La confrontation des trois listes de formules chimiques ainsi obtenues a ensuite mis en évidence l'unique formule

commune correspondant à la formule du métabolite secondaire fongique ionisé (exemple en Tableau 4).

$m/z$ $^{12}\text{C}$ 553,2814	$m/z$ $^{13}\text{C}$ 585,3857	$m/z$ $^{13}\text{C}^{15}\text{N}$ 569,3302
8 formules possibles	11 formules possibles	>400 formules possibles
<b>Calcul du nombre de carbones</b> $N_{\text{C}} = 32$		
2 formules possibles $\text{C}_{32}\text{H}_{34}\text{N}_8\text{Na}$ $\text{C}_{32}\text{H}_{41}\text{O}_8$		>400 formules possibles
<b>Calcul du nombre d'azotes</b> $N_{\text{N}} = 0$		
1 formule possible $\text{C}_{32}\text{H}_{41}\text{O}_8$		

Tableau 4. Exemple du calcul de la composition en carbone et en azote de l'ion détecté à  $m/z_{12\text{C}}=553,2814$  et détermination de sa formule brute. Les formules chimiques ont été générées par le logiciel Xcalibur (Thermo Scientific) selon une précision de masse de 5 ppm et en acceptant les compositions chimiques suivantes : carbone ( $^{12}\text{C}$  et/ou  $^{13}\text{C}$ ) (entre 0 et 60 atomes), hydrogène (entre 0 et 80 atomes), azote ( $^{14}\text{N}$  ou  $^{15}\text{N}$ ) (entre 0 et 25 atomes), oxygène (entre 0 et 25 atomes), sodium (entre 0 et 2 atomes), phosphore (entre 0 et 9 atomes), soufre (entre 0 et 10 atomes), fluor (entre 0 et 16 atomes), chlore (entre 0 et 10 atomes), brome (entre 0 et 5 atomes) et silicium (entre 0 et 8 atomes).

Afin de faciliter cette étape de détermination des formules brutes, le logiciel MassToFormulaCompare a été précédemment développé au sein du laboratoire. Il permet le calcul des formules possibles à partir des trois masses exactes renseignées par l'utilisateur (Figure 20A) en considérant le marquage isotopique et les compositions atomiques minimales et maximales suggérées (Figure 20B). La confrontation de ces trois listes permet de définir l'unique formule commune et donc de connaître la composition chimique de la molécule étudiée (Figure 20C).

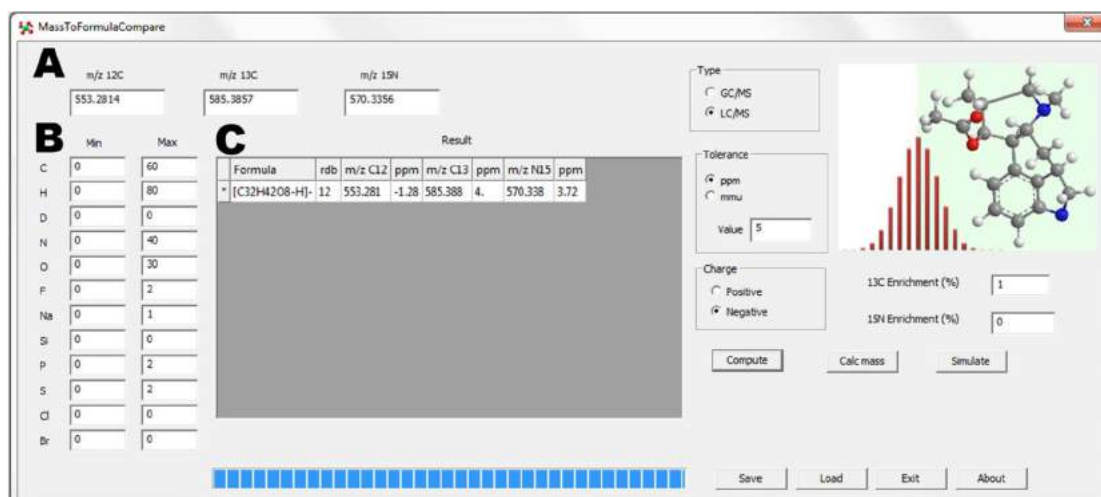


Figure 20. Utilisation du logiciel MassToFormulaCompare pour la détermination de la formule chimique d'un composé à partir du  $m/z$  de l'ion non marqué (ici 511,2909), de l'ion marqué  $^{13}\text{C}$  (ici 539,3837) et de l'ion doublement marqué  $^{13}\text{C}^{15}\text{N}$  (ici 529,3242). A. Rapports  $m/z$  renseignés par l'utilisateur. B. Compositions chimiques minimales et maximales suggérées par l'utilisateur. C. Formule proposée par le logiciel.

Les formules chimiques générées par MassToFormulaCompare sont soumises aux sept règles d'or émises par Kind et Fiehn<sup>2</sup>. Ces règles sont fondées sur la répartition des éléments chimiques dans la nature et la composition générale des molécules naturelles. Elles permettent ainsi de restreindre le nombre de formules chimiques possibles pour une masse exacte observée en spectrométrie de masse :

(1) *Restriction du nombre d'éléments*. Seuls les éléments constituant les molécules naturelles ou résultant des conditions analytiques utilisées au sein du laboratoire sont considérés. Les quantités maximales pour chaque type d'atomes sont limitées par la quantité maximale observée dans les bases de données de produits naturels (Dictionary of Natural Products et Wiley mass spectral). Ainsi, une molécule de 500 Da ne pourra pas comporter plus de 39 atomes de carbones, 72 atomes d'hydrogènes, 20 atomes d'azote, 20 atomes d'oxygène, 9 atomes de phosphore, 10 atomes de soufre, 16 atomes de fluor, 10 atomes de chlore, 5 atomes de brome et 8 atomes de silicium.

(2) *Règles chimiques de Lewis et Senior*. Elles prennent en considération la valence des différents éléments et leur ionisation en spectrométrie de masse.

(3) *Respect des ratios isotopiques*. Les molécules naturelles contiennent les éléments chimiques sous différentes formes isotopiques dont le ratio est déterminé par l'abondance naturelle de chacun des isotopes. Ce ratio influence la masse exacte du composé. Cette règle n'a pas d'impact pour les éléments monoisotopiques tels que le sodium, le fluor et le phosphore.

(4) *Ratio hydrogène/carbone*. L'observation des bases de données de molécules naturelles a montré que 99,7% de ces composés ont un ratio H/C compris entre 0,2 et 3. De rares exceptions sont observées telles que la méthylhydrazine ( $\text{CH}_6\text{N}_2$  H/C=6) ou le tetracyanopyrrole ( $\text{C}_8\text{HN}_5$  H/C= 0,125).

(5) *Ratio hétéroatome/carbone*. L'observation des bases de données de produits naturels a démontré que les ratios de quantités d'azote, d'oxygène, de phosphore et de soufre *versus* de carbone sont compris entre 0 et 1,5. En particulier, l'oxygène et l'azote, qui sont les deux hétéroatomes les plus rencontrés dans la composition des métabolites secondaires fongiques, ont des ratios maximums respectifs de 1,2 et 1,3.

(6) *Probabilités de ratios d'éléments*. Cette règle permet de considérer la composition globale d'un composé au-delà de la validation des règles 4 et 5 vérifiant les ratios de chaque élément chimique de manière distincte. En effet, la formule  $\text{C}_{26}\text{H}_{28}\text{N}_{17}\text{O}_1\text{P}_3\text{S}_8$  valide les précédentes règles, mais la composition globale est improbable pour une molécule naturelle.

(7) *Présence de composés triméthylsilylés*. Le groupement triméthylsilyl (TMS), composé de trois groupements méthyl liés à un atome de silicium [ $-\text{Si}(\text{CH}_3)_3$ ], est un groupement utilisé en chromatographie en phase gazeuse couplée à la spectrométrie de masse pour dériver les molécules à analyser. Cette règle n'impacte donc pas les analyses effectuées en LC/MS dans le cadre de cette thèse.

Les données issues des étapes d'extraction des pics, de suppression de la redondance de l'information et de détermination des formules lors de l'analyse du métabolome secondaire de *P. verrucosum* sont présentées dans l'article 2 en fin de chapitre.

### 1.3. Analyse des métabolites secondaires par MS/MS

Dans le but de poursuivre l'analyse structurale des métabolites secondaires détectés, au-delà de la caractérisation de leur formule chimique, les spectres MS/MS de chacun de ces composés ont été générés. Trois différents types d'analyses MS/MS ont été réalisés.

#### 1.3.1. Génération des spectres MS/MS par acquisition data dépendante

La fragmentation par acquisition data dépendante (DDA) est une dissociation induite par collision (CID, voir section 1.3.2) réalisée de manière automatique sur les  $n$  ions les plus intenses détectés par l'analyseur dans un scan HRMS donné ( $n$  étant fixé par l'opérateur). Dans le cadre de cette thèse, elle a pour objectif d'obtenir des informations structurales dès la première analyse des échantillons (blé  $^{13}\text{C}^{15}\text{N}$  infesté, blé  $^{13}\text{C}$  infesté, blé  $^{12}\text{C}$  infesté et témoin). Ce type d'acquisition peut notamment assister l'étape 3 de l'extraction des pics (Figure 19). De manière à ne pas accumuler les analyses, ce qui provoquerait une baisse de la résolution de chacune d'entre elles, ces fragmentations automatiques ont été détectées en basse résolution dans le piège à ions linéaire (LTQ, Figure 15) et sur les trois ions majoritaires de chaque scan HRMS.

Néanmoins, la fragmentation en mode DDA n'est pas appliquée spécifiquement aux métabolites secondaires du champignon. De plus, il arrive qu'un ion compte parmi les trois ions majoritaires pendant un court temps d'analyse. Son spectre de fragmentation moyen est dans ce cas limité à un faible nombre de spectres, ce qui peut être insuffisant pour obtenir des spectres moyens de bonne sensibilité. Des analyses ciblées ont donc été réalisées sur chacun des métabolites secondaires fongiques après l'étape d'extraction des pics et de suppression de la redondance de l'information.

#### 1.3.2. Analyse ciblée des métabolites secondaires fongiques en mode CID

Des analyses MS/MS ciblées en haute résolution ont été réalisées afin d'obtenir des informations structurales de qualité sur tous les métabolites secondaires, quelle que soit leur intensité relative. Pour cela, chaque composé fongique a été fragmenté précisément à son temps de rétention. Pour chaque molécule, l'énergie de collision a été adaptée afin d'obtenir des fragmentations optimales (fragmentation de la molécule et présence de l'ion précurseur sur le spectre MS/MS). Dans un premier temps, ces fragmentations ciblées ont été réalisées en mode CID dans la trappe linéaire<sup>3</sup>.

Lors de fragmentations en mode CID dans la trappe linéaire, une fréquence correspondant à la fréquence de rotation des ions isolés dans le piège est appliquée sur deux des pôles (axe  $x$ ) provoquant l'accélération des ions. Ces derniers subissent alors des collisions avec le gaz utilisé pour contrôler la charge d'espace (hélium) et emmagasinent de l'énergie interne jusqu'à fragmentation. Lors de ce mode de fragmentation, les ions n'accumulent l'énergie interne suffisante que pour une seule fragmentation. Il est donc rare que des fragmentations consécutives surviennent. Il est néanmoins possible d'étudier la filiation entre les ions fragments par des expériences de spectrométrie de masse successives ( $\text{MS}^n$ ). Cependant, lorsqu'elle est réalisée dans la trappe linéaire, la fragmentation en mode CID possède une restriction de masse empêchant la détection des ions fragments de  $m/z$  inférieur à environ 30% du  $m/z$  de l'ion précurseur.

### 1.3.3. Analyse ciblée des métabolites secondaires fongiques en mode HCD

Des analyses MS/MS ciblées ont également été réalisées en mode HCD (Higher-energy Collisional Dissociation) afin d'obtenir des informations complémentaires à celles générées en mode CID dans la trappe linéaire.

Le mode HCD est un mode de fragmentation non résonant par dissociations induites par collisions, qui s'effectue dans une cellule de collision de type octopole située après la C-trap<sup>4</sup> (voir Figure 15). La détection des ions fragments après une fragmentation HCD est systématiquement réalisée en haute résolution dans l'Orbitrap. Le mode HCD fait partie des modes de collision de basse énergie. Après isolement dans la C-trap, l'ion précurseur est envoyé dans la cellule de collision HCD remplie de gaz (N<sub>2</sub>). L'ion  $\gamma$  est accéléré à une vitesse à laquelle les collisions avec le gaz sont dissociatives. Contrairement au mode CID réalisé dans la trappe linéaire, l'excitation non résonnante du mode HCD est susceptible de générer des fragmentations consécutives. De plus, le mode HCD ne présente pas de restriction de masse, ce qui permet la détection d'ions fragments de faible  $m/z$ .

Les paramètres utilisés pour les analyses MS/MS et les résultats des fragmentations HCD et CID des métabolites secondaires sont présentés en détail dans les articles 2, 3 et 4.

## 2. INTERPRÉTATION DES RÉSULTATS

---

### 2.1. Les bases de données

La première étape de l'interprétation des résultats consiste en une étape de déréplication, c'est-à-dire d'annotation des formules obtenues lors de l'analyse des spectres de masse. Pour ce faire, les formules sont confrontées à des bases de données. L'importance de l'étape de déréplication et du choix des bases de données pour la recherche de métabolites secondaires fait l'objet d'une section particulière dans l'article 1 présenté en introduction. Dans le cadre de cette étude, les résultats ont été soumis à AntiBase 2012<sup>5</sup>. Cette base de données spécialisée dans les produits naturels recense plus de 40 000 métabolites secondaires (de champignons, de bactérie, d'algues, etc.). Les molécules qu'elle comporte sont référencées par propriétés descriptives (masse, formule), physicochimiques (point de fusion, pouvoir rotatoire), spectroscopiques (UV, IR, RMN, HRMS) et biologiques (activité pharmacologique, toxicité), lorsque celles-ci sont disponibles. Dans le cadre de cette thèse, certaines formules reconnues par AntiBase étaient l'objet de multiples suggestions. Les propositions de structures d'origine fongique et en particulier les molécules produites par des moisissures du genre *Penicillium* ont donc été privilégiées. Pour compléter les données d'AntiBase, des recherches bibliographiques ont été effectuées.

### 2.2. Validation des annotations : identification

En raison de la faible disponibilité de composés de référence des métabolites secondaires fongiques dans le commerce, les annotations proposées ont été confrontées aux spectres de fragmentation expérimentaux. Cette interprétation permet un premier tri des identifications en mettant en évidence des incohérences entre la structure supposée et le spectre MS/MS expérimental (par exemple, une structure hypothétique sans fonction acide carboxylique ou ester en désaccord avec un spectre MS/MS comportant une perte de CH<sub>2</sub>O<sub>2</sub>). Dans un deuxième temps, les hypothèses structurales ont été vérifiées selon les quatre niveaux d'identification décrits par Sumner *et al.*<sup>6</sup>

(1) *Identification de niveau 1, « Composé identifié »*. Métabolite présentant au moins deux données indépendantes identiques à celles d'un composé de référence analysé dans les mêmes conditions expérimentales (temps de rétention et spectre de masse, masse précise et spectre MS/MS etc.). Dans le cadre de cette thèse, les molécules ont été identifiées à un niveau 1 lorsqu'elles présentaient un temps de rétention, un profil MS et un spectre MS/MS identiques à ceux d'un standard.

(2) *Identification de niveau 2, « Composé potentiellement annoté »*. Métabolite présentant au moins une donnée identique à celle d'un composé décrit dans la littérature ou fournie par un laboratoire extérieur, ou interprétation des résultats permettant de donner une structure précise en termes d'isomérisation de position (par RMN par exemple).

(3) *Identification de niveau 3, « Classe du composé potentiellement caractérisée »*. Composé dont le composé standard n'est pas disponible et dont aucune donnée analytique n'est précisée dans la littérature. Un composé est identifié au niveau 3 lorsque sa classe est définie (peptide, lipide etc.) ou lorsque la structure est définie mais sans que l'on puisse préciser de quel isomère il s'agit. Dans le cadre de cette étude, un métabolite a été qualifié comme identifié au niveau 3 lorsque sa masse était en accord avec celle d'un composé connu et que son spectre de fragmentation était compatible avec cette structure.

Afin d'assister cette étape, l'outil informatique de simulation de spectres MS/MS, CFM-ID<sup>7</sup>, a été utilisé. Bien que ses résultats ne constituent pas une preuve d'identité du métabolite, ce logiciel permet de conforter des hypothèses structurales en simulant des spectres MS/MS à partir de structures. Dans le cadre de l'analyse des métabolites de *P. verrucosum*, les hypothèses structurales ont été soumises à CFM-ID pour qu'il en génère les spectres de fragmentation hypothétiques. Ceux-ci ont ensuite été comparés aux spectres expérimentaux.

(4) *Identification de niveau 4, « Composé inconnu »*. Métabolite dont aucune hypothèse structurale n'est formulée mais qui a été détecté et dont les paramètres analytiques (temps de rétention et masse exacte définis) permettent de le différencier.

À l'issue de cette étape d'identification, les analyses ont été concentrées sur l'élucidation structurale de composés inconnus. L'observation seule de la formule chimique des composés inconnus ne permettant pas d'orienter la suite des analyses structurales, une étape supplémentaire de traitement des données a été mise en œuvre par la réalisation de réseaux moléculaires.

### **2.3. Réseaux moléculaires : Le "GNPS Molecular Networking System"**

L'outil informatique disponible en ligne du GNPS (Global Natural Product Social molecular networking)<sup>8</sup> permet d'organiser des données de spectrométrie de masse sous forme de réseaux moléculaires (Figure 21). Par la comparaison des spectres MS/MS vectorisés, il permet d'établir des degrés de similarité entre les profils de fragmentation. Puisque deux molécules qui se fragmentent de manières similaires sont susceptibles d'avoir des structures similaires, ce logiciel permet de lier les métabolites selon leur potentielle similarité structurale.

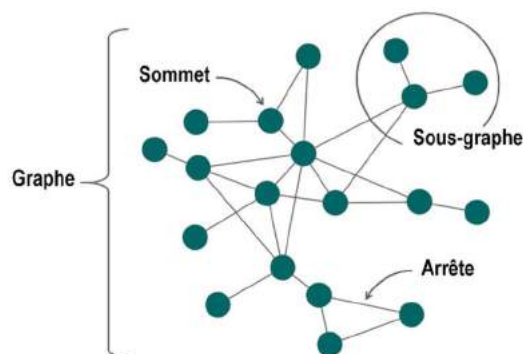


Figure 21. Lexique relatif aux réseaux

Cette organisation des données permet alors de mettre en évidence des composés inconnus semblables à des métabolites secondaires connus et ainsi de faciliter leur élucidation structurale. En se basant sur les mécanismes de fragmentation préalablement interprétés à partir d'une structure connue, l'élucidation d'un spectre semblable est, en effet, plus aisée que son interprétation sans *a priori*. D'autre part, en liant les molécules selon leurs potentielles similarités structurales et donc selon leur classe chimique (alkaloïdes,  $\beta$ -lactames etc.), l'observation d'un graphe ou sous-graphe contenant plusieurs ions inconnus peut permettre la mise en évidence de nouvelles classes de métabolites.

### 3. LE SÉQUENÇAGE PEPTIDIQUE *DE NOVO*

Lors de l'analyse des métabolomes secondaires de *P. verrucosum* et de *P. nordicum*, de nouveaux peptides cycliques et linéaires ont été mis en évidence (Articles 2 et 3). Leurs structures ont été investiguées par séquençage peptidique *de novo* à partir des mécanismes de fragmentation peptidique décrits dans cette section.

#### 3.1. Les outils de séquençage peptidique automatique

Avec l'essor de la protéomique, des outils informatiques d'aide au séquençage peptidique ont été développés. Parmi les programmes disponibles en libre accès ou par licence, on distingue les logiciels de simulation de spectres MS/MS et les logiciels d'annotation de spectres MS/MS. Les premiers sont présentés comme un outil de vérification permettant de s'assurer que le spectre MS/MS d'un peptide donné est en accord avec les fragmentations peptidiques théoriques. Ainsi, l'outil MS-Product de Protein Prospector ([prospector.ucsf.edu](http://prospector.ucsf.edu)) permet de calculer les ions fragments classiques générés à partir d'un peptide donné. Bien que cet outil soit largement utilisé pour vérifier une identification *de novo*, le logiciel, dit conservateur, est basé sur des règles de fragmentation limitées et ne précise pas la probabilité d'apparition des ions. Les logiciels d'annotation de spectres MS/MS sont présentés comme des outils de séquençage peptidique à proprement parlé puisqu'ils permettent de déduire la structure du peptide précurseur à partir d'un spectre inconnu. De nombreux logiciels pour le séquençage de peptide sont basés sur la confrontation des spectres MS/MS inconnus avec ceux référencés dans les bases de données de protéomique. Si le peptide est déjà référencé, la cohérence des spectres permet une identification du peptide étudié. Divers outils de ce type sont disponibles, tels que MASCOT<sup>9</sup>, SEQUEST<sup>10</sup>, X!TANDEM<sup>11</sup> ou OMSSA<sup>12</sup>. Néanmoins, cette méthode ne fonctionne que pour les peptides référencés et son utilisation n'est pas recommandée pour l'élucidation structurale de peptides produits par des organismes reconnus pour leurs fréquentes

modifications post-traductionnelles<sup>13</sup>. De plus, le spectre MS/MS à annoter doit avoir été obtenu à l'aide du même mode de fragmentation que celui ayant généré les spectres des bases de données, afin qu'ils soient comparables.

Quel que soit l'outil utilisé, il est nécessaire de vérifier les annotations suggérées<sup>14</sup>. Ces programmes sont particulièrement utilisés pour faciliter l'analyse de jeux de données très importants générés notamment lors de digestions trypsiniques pour le séquençage de protéiques. De nombreux auteurs attirent l'attention sur le fait que les résultats de ces outils informatiques doivent être considérés avec précaution. Dans leur revue, Medzihradsky *et al.* insistent sur l'importance de l'interprétation manuelle des spectres MS/MS en peptidomique, qui permet d'une part la vérification des résultats bioinformatiques et d'autre part la prise en considération de mécanismes de fragmentation non intégrés dans ces programmes<sup>13</sup>. Pour ces raisons, la vérification des annotations des fungusporines et l'élucidation structurale des potentiels nouveaux peptides sont réalisées manuellement dans cette étude. Pour cela, les mécanismes de fragmentation peptidique opérant après ionisation en électrospray positif et en modes HCD et CID de basse énergie sont détaillés et appliqués à l'étude des composés étudiés. La fragmentation des peptides linéaires est référencée dans de nombreux articles. Les mécanismes détaillés dans les sections suivantes à ce propos s'appuient notamment sur les démonstrations de Paizs *et al.*<sup>15</sup>. Les mécanismes de fragmentation des peptides adduits au sodium et des peptides cycliques sont quant à eux l'objet d'un faible nombre d'études. Les différents points de vue sont développés dans cette partie.

### **3.2. Le séquençage peptidique de novo**

#### **3.2.1. Le modèle du proton mobile**

Il est communément admis que les fragmentations obtenues par dissociation induite par collision de faible énergie peuvent être annotées sur la base du « Mobile Proton Model ». Les collisions de type CID de haute énergie font référence aux énergies de collisions de l'ordre du keV générées notamment dans les appareils de types TOF-TOF ou magnétiques. Au contraire, les activations par basse énergie de collision, telles que celles appliquées lors d'activations de type CID en trappe à ions ou de type HCD, sont de l'ordre des dizaines d'eV. Deux types de peptides peuvent être différenciés : les peptides dans lesquels un groupement va séquestrer la charge et favoriser un site de fragmentation particulier, et ceux chez lesquels le proton peut être capté par plusieurs fonctions permettant aux dissociations de survenir à différents sites de la molécule. En spectrométrie de masse utilisant des modes d'ionisation à pression atmosphérique (API) et des modes de fragmentation de basse énergie, les fragmentations se situent en grande majorité au niveau des liaisons peptidiques. À condition qu'aucune chaîne latérale ne séquestre la charge, la protonation d'un peptide s'effectue de manière plus favorable sur l'azote d'une amine (N-terminale ou latérale). Le modèle du proton mobile stipule que le proton peut ensuite migrer le long des liaisons peptidiques. La dissociation de la molécule est ensuite généralement dirigée par la charge.

#### **3.2.2. Étude théorique des mécanismes de fragmentation des peptides linéaires**

Dans des expériences de séquençage *de novo*, ce sont les fragmentations situées au niveau des liaisons peptidiques qui donnent des informations quant à la séquence du peptide. Selon la zone et l'orientation de ces fragmentations, différentes séries d'ions sont générées selon la nomenclature



proposée par Biemann (1990)<sup>16</sup> (Figure 22). Dans le cas d'une ionisation par électrospray et d'une fragmentation par activation de type CID comme dans le cadre de cette thèse, les dissociations génèrent principalement des ions des séries **y**, **b** et **a**. Les différences de masse entre deux ions consécutifs de la même série représentent les masses des résidus des acides aminés. Le séquençage peptidique *de novo* repose donc sur l'annotation de ces séries d'ions et sur la confirmation de la séquence par une compréhension globale des mécanismes de fragmentation. Les réactions de dissociation à l'origine de ces ions fragments classiques sont présentées dans cette partie selon le modèle d'un tétrapeptide dont aucune des chaînes latérales ne présente de groupements fonctionnels capables de capter le proton ou d'influencer la dissociation. Les acides aminés sont numérotés de l'extrémité N terminale à l'extrémité C terminale.

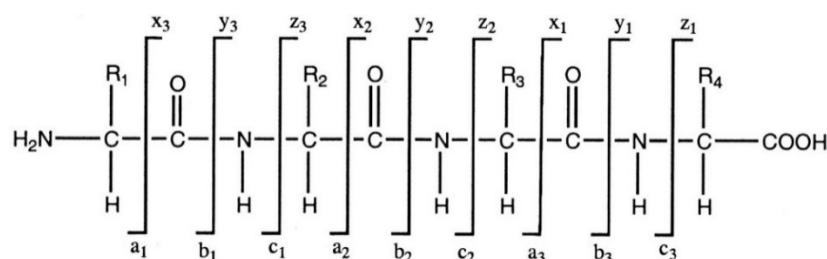


Figure 22. Nomenclature de la fragmentation peptidique en spectrométrie de masse.

### Mécanisme général de pré-dissociation et de dissociation

La protonation d'un tétrapeptide sans groupement fonctionnel latéral se réalise dans un premier temps au niveau de la fonction amine N terminale. Le proton mobile est dans un deuxième temps capté par l'atome d'azote d'une liaison peptidique, ce qui a pour conséquence d'une part la fragilisation de la liaison peptidique, et d'autre part l'augmentation de l'électrophilie du carbone de l'amide qui devient alors une cible préférentielle pour les attaques nucléophiles. La polarisation de la liaison peptidique entraîne ainsi différents types de dissociation qui peuvent résulter d'un clivage direct de la liaison amide ou d'un réarrangement. Ce dernier est favorisé dans le cas d'analyse en basse énergie car le clivage direct de la liaison peptidique nécessite l'accumulation d'une forte énergie par l'ion. De plus, selon les affinités protoniques et la stabilité des fragments, les dissociations induites par collision de basse énergie peuvent générer des ions dipôles au sein desquels un transfert de proton est possible (exemple Figure 23, groupes entre crochets). Enfin, en HCD, des mécanismes de dissociation consécutives sont possibles en raison de l'excès d'énergie interne communiqué par ce mode CID non résonnant.

### Réarrangement par formation d'une oxazolone et dissociations consécutives

Lors de la protonation de l'atome d'azote d'une liaison peptidique, l'un des mécanismes majeurs induisant la dissociation de la molécule résulte de l'attaque nucléophile de l'atome de carbone de la fonction amide par l'atome d'oxygène de la liaison peptidique voisine côté N terminal (Figure 23 ion **b**). L'azote de la fonction amine terminale peut également réaliser cette attaque nucléophile mais de façon minoritaire. La formation d'une oxazolone protonée induit la rupture de la liaison amide et la formation d'un ion dipôle entre l'oxazolone protonée et le résidu neutre de l'extrémité C terminale du peptide précurseur (voie **b<sub>x</sub>-y<sub>x</sub>**). Le transfert de proton au sein de cet ion dipôle est à l'origine d'ions **y** (Figure 23 ion **y**) en fonction de l'affinité protonique de chacune des parties.

En HCD, l'oxazolone protonée des ions **b** peut s'ouvrir et générer un ion de type **a** par perte de CO (voie  $b_x \rightarrow a_x$ ). Ce processus est le principal mécanisme générant les ions de type **a** à partir desquels pourront être produits également des ions immonium internes notés **I** à partir d'un ion dipôle (voie  $a_x \rightarrow a_{x-1}$ ). Des phénomènes de dissociation consécutive directement à partir d'ion  $b_x$  (voie  $b_x \rightarrow b_{x-1}$ ) sont également observés pour former un ion  $b_{x-1}$ . En HCD, ces dissociations consécutives des ions **b** de haut grade expliquent en partie l'intensité relative plus importante des ions **y**. En CID au contraire, l'activation résonnante générée dans la trappe d'ions ne génère pas de fragmentations consécutives.

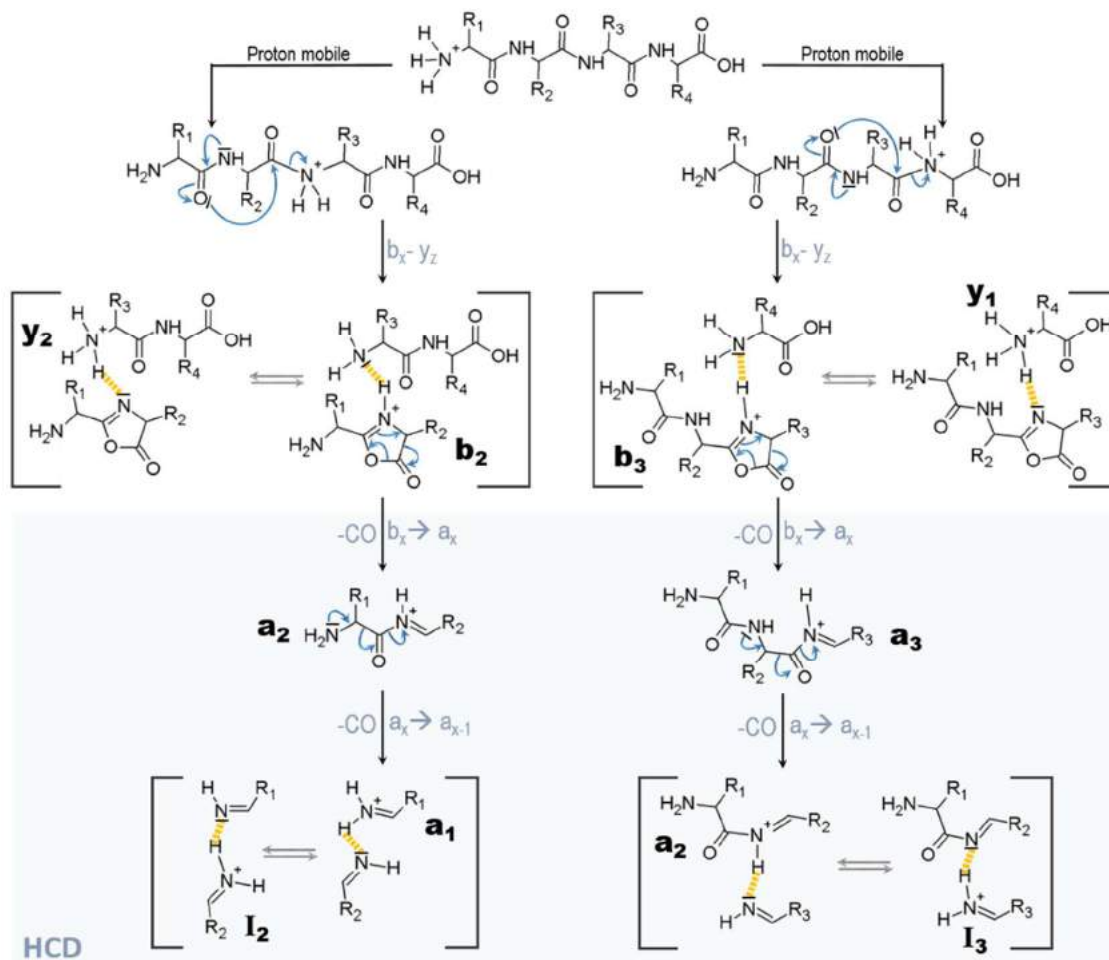


Figure 23. Fragmentation peptidique par formation d'une oxazolone et dissociations consécutives. Les groupements  $R_1, R_2, R_3$  et  $R_4$  représentent les chaînes latérales des acides aminés.

### Réarrangement par formation d'une dicétopipérazine

La formation des ions **y** peut également impliquer la formation d'une dicétopipérazine (Figure 24). La faible affinité protonique de celle-ci entraîne systématiquement une conservation de la charge par le peptide linéaire qui en résulte sous forme d'un ion de type **y**. Ce mécanisme peut être également réalisé par l'attaque de la fonction amine terminale sur une liaison peptidique plus lointaine mais les facteurs entropiques du peptide cyclique généré défavorisent ce type de dissociation.

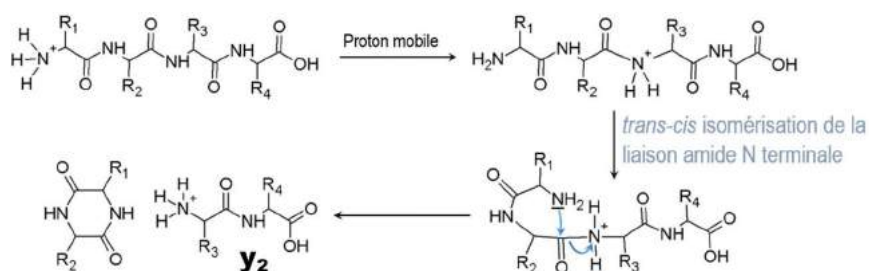


Figure 24. Fragmentation peptidique générée par formation d'une dicétopipérazine.

### Formation des ions *y* - dissociation directe ou réarrangement

Alors que les mécanismes de dissociation décrits précédemment sont favorisés par des réarrangements, une rupture hétérolytique de la liaison peptidique de l'acide aminé côté N terminal (voie  $a_1$ - $y_x$ ) est possible. Dans ce cas, une perte consécutive de monoxyde de carbone est possible et l'affinité protonique de chacune des deux parties de l'ion dipôle qui en résulte détermine alors l'abondance relative d'ions *a* et d'ions *y* formés (Figure 25, gauche).

La formation des ions *y* peut également être issue d'un réarrangement via la formation d'une aziridinone (Figure 25, droite). Celui-ci n'est pas favorisé par rapport à une rupture hétérolytique en raison de la faible stabilité de ce cycle à trois sommets. Le cycle aziridinone protoné sur son azote est stable au sein de l'ion dipôle qu'il forme avec l'extrémité C terminale du peptide précurseur. En revanche, dans le cas où la charge est conservée par l'aziridinone protonée, celle-ci se décompose en un ion *a* par une perte de CO. Des ions *y* sont également générés avec des abondances relatives respectant l'affinité protonique relative des deux parties du dimère.

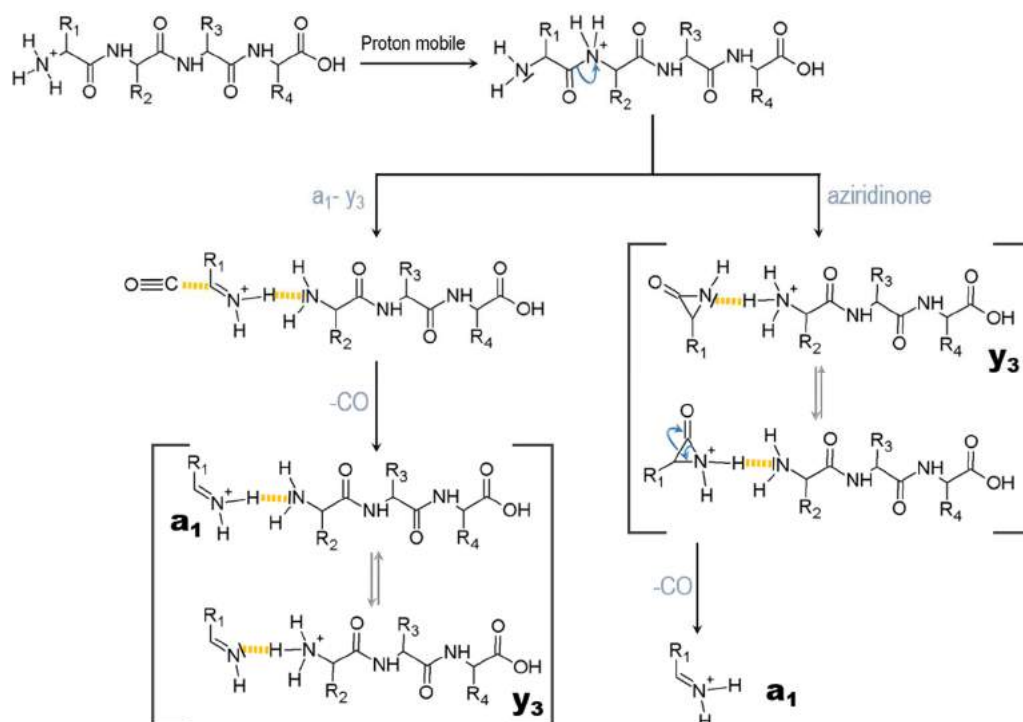


Figure 25. Formation des ions *y* par rupture hétérolytique ou par formation d'une aziridinone

### Fragmentations non informatives quant à la séquence peptidique

Lors de l'analyse de peptides par spectrométrie de masse, certaines dissociations ne sont pas informatives quant à la séquence peptidique. La perte d'eau peut survenir à partir du peptide précurseur ou d'ions fragments par captation du proton mobile par un oxygène de la fonction acide carboxylique terminale. La polarisation de la liaison entraîne sa rupture, suivie par la formation d'un cycle oxazolone (Figure 26). Des pertes d'eau peuvent également survenir au niveau des chaînes latérales de résidus Asp, Glu, Ser et Thr.

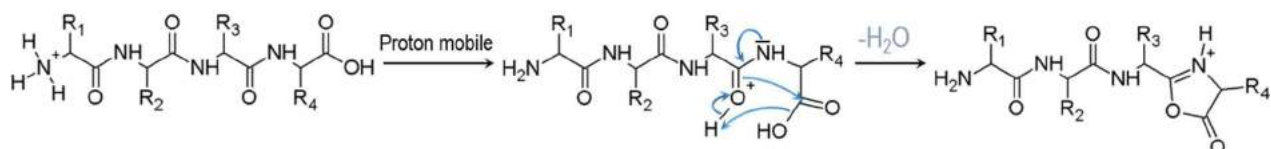


Figure 26. Perte d'eau à partir de la fonction acide carboxylique terminale

La présence d'ions immonium dans la zone des faibles ratios  $m/z$  du spectre MS/MS d'un peptide permet d'émettre des hypothèses quant à la constitution du peptide (Tableau 5) mais n'est pas représentative de la séquence. Seul l'ion immonium correspondant à l'ion issu du premier acide aminé de la séquence ( $a_1$ ) est informatif. La formation des immoniums internes implique le clivage de deux liaisons amides. Le premier survient lors de la génération des ions de types **b** et **y**. Le mécanisme de fragmentation des ions **b** en ions immoniums est présenté Figure 23. Des ions immoniums peuvent également être générés à partir des ions **y** par les voies  $I_x-y_z$  et  $I_x \rightarrow I_{x-1}$  (Figure 27).

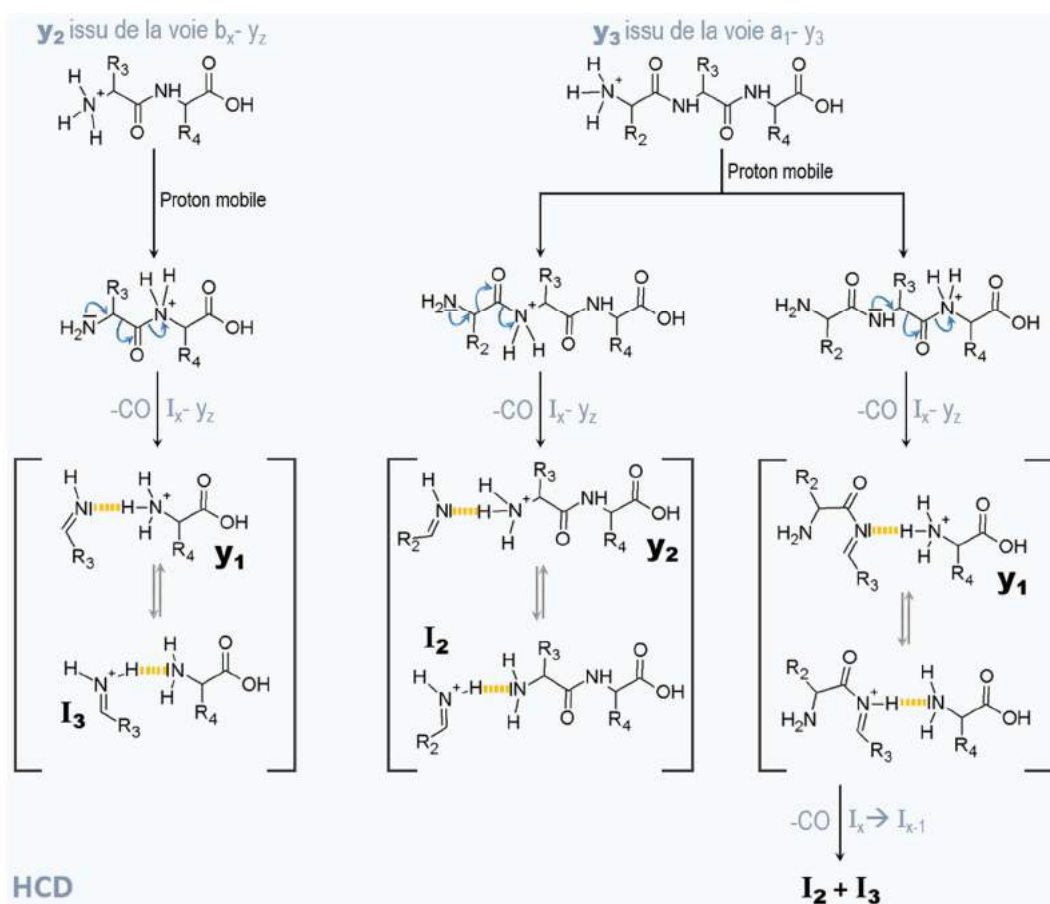


Figure 27. Formation d'ions immonium internes à partir d'ions **y**

Acide aminé	Code à 3 lettres	Code à 1 lettre	Masse du résidu	<i>m/z</i> de l'ion immonium
Alanine	Ala	A	71.0372	44.0491
Arginine	Arg	R	156.0981	129.1131
Asparagine	Asn	N	114.0429	87.0549
Acide aspartique	Asp	D	115.0269	88.0389
Cystéine	Cys	C	103.0092	76.0211
Acide glutamique	Glu	E	129.0426	102.0546
Glutamine	Gln	Q	128.0586	101.0705
Glycine	Gly	G	57.0215	30.0334
Histidine	His	H	137.0589	110.0709
Isoleucine	Ile	I	113.0841	86.0960
Leucine	Leu	L	113.0841	86.0960
Lysine	Lys	K	128.0949	101.1069
Méthionine	Met	M	131.0405	104.0524
Phénylalanine	Phe	F	147.0684	120.0804
Proline	Pro	P	97.0528	70.0647
Sérine	Ser	S	87.0320	60.0439
Thréonine	Thr	T	101.0477	74.0596
Tryptophane	Trp	W	186.0793	159.0913
Tyrosine	Tyr	Y	163.0633	136.0753
Valine	Val	V	99.0684	72.0804

Tableau 5. Informations relatives aux 20 acides aminés protéinogènes standards

### Fragmentation orientée par la nature du résidu

La perte de  $\text{NH}_3$  survient principalement lors de la fragmentation des chaînes latérales des résidus Arginine, Lysine, Asparagine et Glycine. Néanmoins, certains articles rapportent la perte de l'amine terminale en mode CID de basse énergie. Dans ce cas, la présence d'un acide aminé aromatique en position N terminale est nécessaire. Cette dissociation est particulièrement facilitée en présence de tryptophane grâce à l'atome d'azote du noyau indole qui facilite l'isomérisation<sup>17</sup> et l'attaque nucléophile sur l'atome de carbone en position  $\alpha$ . Différents résultats de cyclisation ont été publiés (Figure 28A-D). Des expériences de modélisation moléculaire ont permis de démontrer que la perte d'ammoniac résulte de la création d'un spirocyclopropane<sup>18</sup> (Figure 28B).

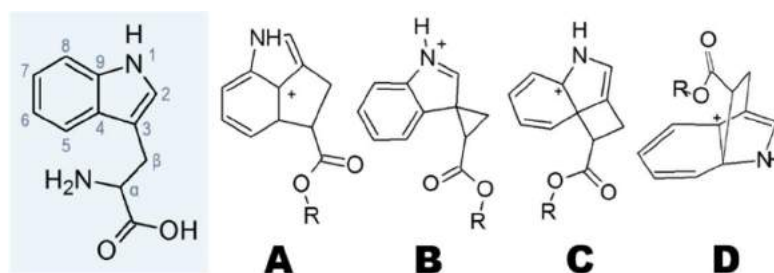


Figure 28. Hypothèses d'attaque nucléophile expliquant la perte de  $\text{NH}_3$  à partir de la fonction amine terminale d'un tryptophane.

### 3.2.3. Les mécanismes de fragmentation des adduits sodium

En spectrométrie de masse en mode CID de basse énergie, les mécanismes de fragmentation des peptides sous forme d'adduits au sodium (ou autres ions pseudo-moléculaires cationisés) sont

différents de ceux des peptides protonés. Renner et Spiteller suggèrent que l'ion sodium interagisse avec la fonction acide carboxylique terminale du peptide linéaire sous forme de zwitterion<sup>19</sup> (Figure 29). Ceci induit la polarisation de la fonction peptidique avec l'acide aminé adjacent et favorise l'attaque nucléophile de son atome de carbone par l'un des atomes d'oxygène de la fonction carboxylate. Le peptide peut alors perdre l'acide aminé C terminal grâce à un réarrangement entraînant une perte de CO et d'une imine<sup>20</sup>. Puisque l'ion précurseur  $[M+Na]^+$  a une composition de type  $[b_x+OH+Na]^+$ , les ions fils résultants sont notés dans la littérature  $[b_{x-1}+OH+Na]^+$ <sup>21</sup>.

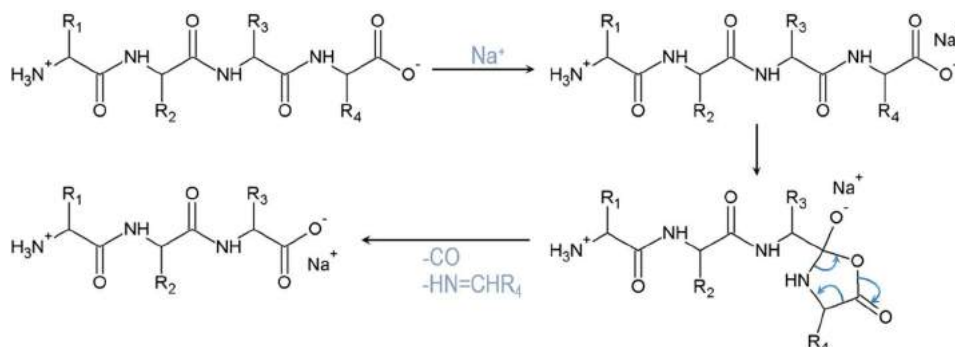


Figure 29. Mécanisme de fragmentation peptidique des adduits sodium d'après Renner et al.<sup>19</sup>

Les spectres MS/MS issus de la fragmentation d'adduits sodium dépendent des acides aminés C terminaux. D'après Lin *et al.*, ces derniers peuvent être distingués selon deux classes. La classe I comporte 15 acides aminés<sup>21</sup>. Parmi eux, 10 acides aminés génèrent exclusivement des ions de type  $[b_{x-1}+OH+Na]^+$ . L'alanine, la valine, la leucine, l'isoleucine, la méthionine, la proline, la phénylalanine, la tyrosine, la lysine, et l'histidine forment ainsi la sous-classe « la ». Pour des chaînes peptidiques de plus de trois acides aminés, la glycine et l'acide aspartique induisent également des fragmentations de classe « la ». La cystéine, l'arginine et l'acide glutamique forment la sous-classe « lb » en raison de fragmentations supplémentaires qu'ils peuvent générer. Le tryptophane, la sérine, la thréonine, l'asparagine et la glutamine forment quant à eux la classe II.

Pour la classe « la », l'ion  $[b_{x-1}+OH+Na]^+$  est l'unique ion généré en mode CID de basse énergie<sup>21</sup>. Les peptides de la classe II peuvent quant à eux générer également des ions décrits précédemment dans le cas de fragmentation de peptides protonés. Ces fragmentations peuvent devenir majoritaires dans le cas de chaînes peptidiques courtes (environ trois acides aminés), ce qui suggère une interaction intramoléculaire entre les extrémités N terminale et C terminale. Fernandez Ocana *et al.* indiquent que des ions de type  $[b_x+Na-H]^+$ ,  $[b_{x-1}+Na-H]^+$  et  $[a_{x-1}+Na-H]^+$  peuvent également être détectés en fonction de la nature de l'acide aminé C terminal<sup>22</sup>. Des ions de type  $[y+Na-H]^+$  ont également été observés<sup>20,22</sup>. La captation du sodium par les différentes fonctions amides du peptide est un processus suggéré par Lee *et al.* et qui pourrait être à l'origine de ce type de dissociation<sup>20</sup>. En effet, des études de modélisation moléculaire ont démontré que l'atome de sodium était maintenu en équilibre au sein d'un térapeptide par interaction avec plusieurs fonctions carboxyles (exemple du peptide AAAA en Figure 30). De plus, la liaison au sodium par une fonction carboxyle est favorisée grâce au doublet de l'azote en position  $\alpha$  du carboxyle<sup>23</sup>.

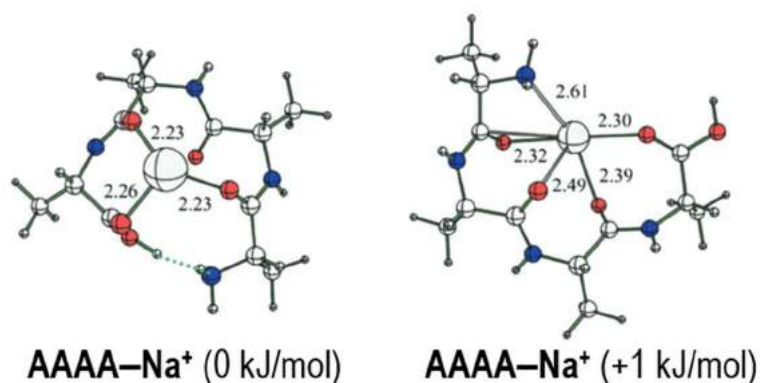


Figure 30. Interaction d'un atome de sodium avec les diverses fonctions d'un tétrapeptide (exemple de AAAA). Extrait de Wang et al.<sup>24</sup> Les chiffres accolés aux arrêtes représentent les distances en angström. Les chiffres entre parenthèse sont les énergies relatives.

Enfin, d'après Lin *et al.*, le tryptophane est le seul acide aminé à ne pas suivre de schéma de dissociation prédictible lorsqu'il est présent à l'extrémité C terminale sous forme d'adduit sodium<sup>21</sup>. Selon la longueur et la séquence du peptide, des ions fragments variés ont été observés, avec des ions fragments majoritaires correspondant à des ions  $[b_{x-1}+OH+Na]^+$  ou à des pertes d'eau issues de l'ion précurseur.

### 3.2.4. Les mécanismes de fragmentation des peptides cycliques

Alors que le séquençage des peptides linéaires peut se faire en une seule étape de spectrométrie de masse, l'élucidation structurale des peptides cycliques est plus complexe et nécessite généralement plusieurs stades de fragmentation<sup>25</sup>. La complexité des spectres MS/MS de peptides cycliques provient de la multitude des sites possibles d'ouverture du cycle et des séries d'ions fragments qui en résultent. L'annotation des ions fragments est soumise à une nomenclature spécifique de type «  $b_{x|z}$  », où J et Z représentent les deux acides aminés (code à une lettre) de part et d'autre de la liaison peptidique ayant permis l'ouverture du cycle (J étant en position N terminale et Z en position C terminale).

Selon le modèle du proton mobile, la charge peut être portée par différentes fonctions chimiques du peptide. En raison de l'absence de fonction amine terminale, le peptide cyclique est protoné au niveau des liaisons peptidiques ou sur les chaînes latérales qui possèdent une fonction basique. La dissociation des liaisons peptidiques se réalise alors selon le même mécanisme que celui des peptides linéaires via une fragmentation dirigée par la charge<sup>15</sup> (Figure 31, voie 2). Après ouverture du cycle, le peptide possède une liaison amine terminale et une oxazolone C terminale, de manière similaire aux ions de type **b** issus des peptides linéaires. Les fragmentations classiques des ions **b** peuvent donc être observées. Notamment, la dissociation de l'oxazolone génère la perte d'un résidu sous forme aziridinone par l'attaque nucléophile de l'oxygène de la fonction amide voisine (Figure 31, voie 4). Ce processus de type  $b_x \rightarrow b_{x-1}$  peut ainsi entraîner la perte successive de chaque résidu depuis l'extrémité C terminale. Lorsque l'attaque nucléophile est générée par l'atome d'oxygène de la fonction amide qui n'est pas adjacente à l'oxazolone, un ion dipôle peut former des ions de type **y** ou **b** par transfert de proton (Figure 31, voie 6). D'après Williams *et al.*, la formation d'ions **y** peut également résulter de la perte de l'acide aminé N terminal sous forme d'aziridinone<sup>26</sup> (Figure 31, voie 5). Enfin, des ions de types **a** sont générés par perte de monoxyde de carbone selon la voie  $b_x \rightarrow a_x$ <sup>15</sup>

(Figure 31, voie 3) ou directement au moment de l'ouverture du cycle (Figure 31, voie 1) par génération d'un ion acylium<sup>27</sup>. Considérant l'instabilité des ions acylium, il est probable qu'une perte concertée de CO s'effectue et que des ions de type **a** soient produits (Figure 31, voie 1). Cette perte de CO n'étant pas explicitement considérée par les auteurs, le mécanisme est indiqué en pointillé sur la Figure 31. Bleiholder *et al.* rapportent une perte d'imine à partir de ces ions **a** selon la voie  $a_x \rightarrow b_{x-1}$ <sup>28</sup>.

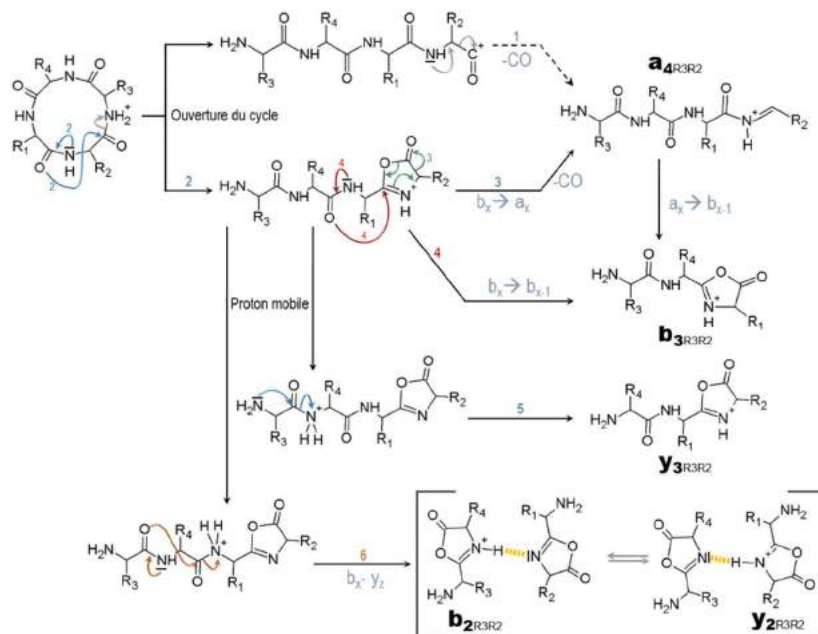


Figure 31. Mécanismes de fragmentation des peptides cycliques protonés en mode CID de basse énergie

Les peptides cycliques sont particulièrement sujets à la formation d'adduits sodium. En effet, la chélation est favorisée au sein du cycle avec les différentes fonctions carboxyles des liaisons peptidiques<sup>29</sup>. Néanmoins, peu d'informations sont disponibles quant aux mécanismes de fragmentation mis en œuvre en spectrométrie de masse. Suite à l'étude de la fragmentation d'un cyclopentapeptide, Ngoka *et al.* ont suggéré que l'ion sodium soit complexé par une fonction carboxyle<sup>29</sup>. L'ouverture du cycle serait donc générée à distance du site de fixation du sodium, par formation d'une aziridinone en position C terminale. Des pertes consécutives d'acides aminés sous forme d'aziridinones peuvent ainsi se poursuivre depuis les extrémités C terminales (Figure 32, voie 2) et N terminales. À chaque étape de formation d'une aziridinone, la perte de CO peut mener à une forme imine (Figure 32, voie 1).

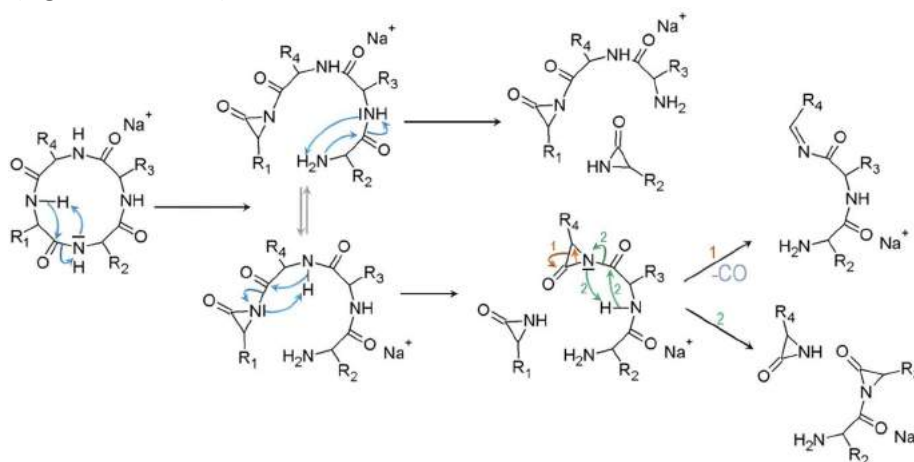


Figure 32. Mécanismes de fragmentation des peptidiques cycliques adduits au sodium d'après Ngoka *et al.*<sup>1</sup>



Outre le mécanisme proposé par Ngoka *et al.*, il est possible que la complexation du sodium induise la formation d'un zwitterion. Ainsi, les fragmentations seraient dirigées par la charge (Figure 33).

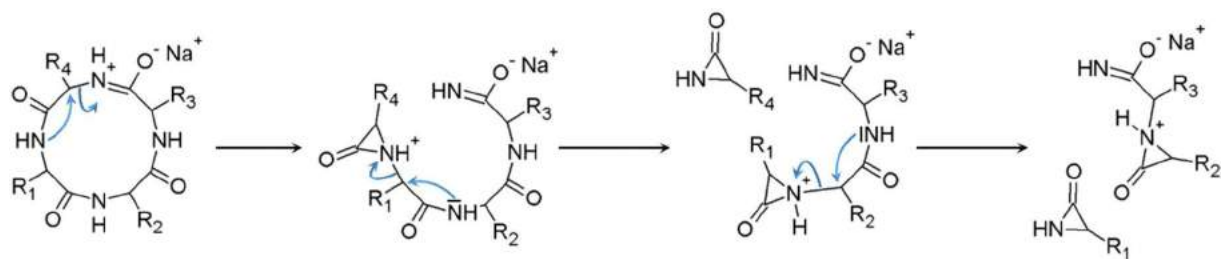
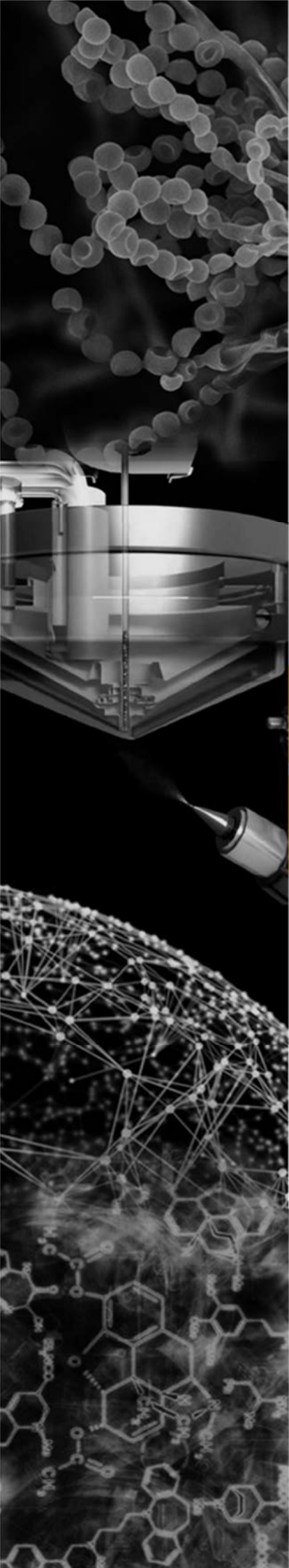


Figure 33. Complexation du sodium sous forme de zwitterion et fragmentation des peptidiques cycliques adduits au sodium.

## 4. RÉFÉRENCES DU CHAPITRE MATÉRIELS &amp; MÉTHODES

- 1 P. M. Cano, E. L. Jamin, S. Tadrist, P. Bourdaud'Hui, M. Péan, L. Debrauwer, I. P. Oswald, M. Delaforge and O. Puel, *Anal. Chem.*, 2013, **85**, 8412–8420.
- 2 T. Kind and O. Fiehn, *BMC Bioinformatics*, 2007, **8**, 105.
- 3 M. Pejchinovski, J. Klein, A. Ramirez-Torres, V. Bitsika, G. Mermelekas, A. Vlahou, W. Mullen, H. Mischak and V. Jankowski, *Proteomics - Clin. Appl.*, 2015, **9**, 531–542.
- 4 J. V Olsen, B. Macek, O. Lange, A. Makarov, S. Horning and M. Mann, *Nat. Methods*, 2007, **4**, 709–12.
- 5 H. Laatsch, 2012.
- 6 L. W. Sumner, A. Amberg, D. Barrett, M. H. Beale, R. Beger, C. A. Daykin, T. W. Fan, O. Fiehn, R. Goodacre, J. L. Griffin, T. Hankemeier, N. Hardy and J. Harnly, *NIH*, 2007, **3**, 211–221.
- 7 F. Allen, A. Pon, M. Wilson, R. Greiner and D. Wishart, *Nucleic Acids Res.*, 2014, **42**, 1–6.
- 8 M. Wang, J. J. Carver, V. V Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. Kapon, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A. V Melnik, M. J. Meehan, W.-T. Liu, M. Crüsemann, P. D. Boudreau, E. Esquenazi, M. Sandoval-Calderón, R. D. Kersten, L. A. Pace, R. A. Quinn, K. R. Duncan, C.-C. Hsu, D. J. Floros, R. G. Gavilan, K. Kleigrewe, T. Northen, R. J. Dutton, D. Parrot, E. E. Carlson, B. Aigle, C. F. Michelsen, L. Jelsbak, C. Sohlenkamp, P. Pevzner, A. Edlund, J. McLean, J. Piel, B. T. Murphy, L. Gerwick, C.-C. Liaw, Y.-L. Yang, H.-U. Humpff, M. Maansson, R. A. Keyzers, A. C. Sims, A. R. Johnson, A. M. Sidebottom, B. E. Sedio, A. Klitgaard, C. B. Larson, C. A. Boya P, D. Torres-Mendoza, D. J. Gonzalez, D. B. Silva, L. M. Marques, D. P. Demarque, E. Pociute, E. C. O'Neill, E. Briand, E. J. N. Helfrich, E. A. Granatosky, E. Glukhov, F. Ryffel, H. Houson, H. Mohimani, J. J. Kharbush, Y. Zeng, J. A. Vorholt, K. L. Kurita, P. Charusanti, K. L. McPhail, K. F. Nielsen, L. Vuong, M. Elfeki, M. F. Traxler, N. Engene, N. Koyama, O. B. Vining, R. Baric, R. R. Silva, S. J. Mascuch, S. Tomasi, S. Jenkins, V. Macherla, T. Hoffman, V. Agarwal, P. G. Williams, J. Dai, R. Neupane, J. Gurr, A. M. C. Rodríguez, A. Lamsa, C. Zhang, K. Dorrestein, B. M. Duggan, J. Almaliti, P.-M. Allard, P. Phapale, L.-F. Nothias, T. Alexandrov, M. Litaudon, J.-L. Wolfender, J. E. Kyle, T. O. Metz, T. Peryea, D.-T. Nguyen, D. VanLeer, P. Shinn, A. Jadhav, R. Müller, K. M. Waters, W. Shi, X. Liu, L. Zhang, R. Knight, P. R. Jensen, B. Ø. Palsson, K. Pogliano, R. G. Linington, M. Gutiérrez, N. P. Lopes, W. H. Gerwick, B. S. Moore, P. C. Dorrestein and N. Bandeira, *Nat. Biotechnol.*, 2016, **34**, 828–837.
- 9 D. N. Perkins, D. J. C. Pappin, D. M. Creasy and J. S. Cottrell, *Electrophoresis*, 1999, **20**, 3551–3567.
- 10 J. K. Eng, A. L. McCormack and J. R. Yates, *Am. Soc. Mass Spectrom.*, 1994, **5**, 976–989.
- 11 R. Craig and R. C. Beavis, *Bioinformatics*, 2004, **20**, 1466–1467.
- 12 L. Y. Geer, S. P. Markey, J. a Kowalak, L. Wagner, M. Xu, D. M. Maynard, X. Yang, W. Shi and S. H. Bryant, *J. Proteome Res.*, 2004, **3**, 958–964.
- 13 K. F. Medzihradsky and R. J. Chalkley, *Mass Spectrom. Rev.*, 2015, **34**, 43–63.
- 14 J. Seidler, N. Zinn, M. E. Boehm and W. D. Lehmann, *Proteomics*, 2010, **10**, 634–649.
- 15 B. Paizs and S. Suhail, *Mass Spectrom. Rev.*, 2005, **24**, 508–548.
- 16 K. Biemann, *Methods Enzymol.*, 1990, **193**, 886–887.
- 17 H. El Aribi, G. Orlova, A. C. Hopkinson and K. W. M. Siu, *J. Phys. Chem. A*, 2004, **108**, 3844–3853.
- 18 H. Lioe, R. a J. O'Hair and G. E. Reid, *J. Am. Soc. Mass Spectrom.*, 2004, **15**, 65–76.
- 19 D. Renner and G. Spiteller, *Biomed. Environ. Mass Spectrom.*, 1988, **15**, 75–77.
- 20 S. W. Lee, H. S. Kim and J. L. Beauchamp, *J. Am. Chem. Soc.*, 1998, **120**, 3188–3195.
- 21 T. Lin, A. H. Payne and G. L. Glish, *J. Am. Soc. Mass Spectrom.*, 2001, **12**, 497–504.
- 22 M. F. Ocaña, J. Jarvis, R. Parker, P. M. Bramley, J. M. Halket, R. K. P. Patel and H. Neubert, *Proteomics*, 2005, **5**, 1209–1216.
- 23 B. A. Cerda, S. Hoyau, G. Ohanessian and C. Wesdemiotis, *J. Am. Chem. Soc.*, 1998, **120**, 2437–3448.
- 24 P. Wang, C. Wesdemiotis, C. Kapota and G. Ohanessian, *J. Am. Soc. Mass Spectrom.*, 2007, **18**, 541–552.
- 25 C. Jia, W. Qi, Z. He and B. Qiao, *Cent. Eur. J. Chem.*, 2006, **4**, 620–631.
- 26 S. M. Williams and J. S. Brodbelt, *J Am Soc Mass Spectrom*, 2004, **15**, 1039–1054.
- 27 L. C. M. Ngoka and M. L. Gross, *Am. Soc. Mass Spectrometry*, 1999, **305**, 6–9.
- 28 C. Bleiholder, S. Osburn, T. D. Williams, S. Suhail, M. Van Stipdonk, A. G. Harrison and B. Paizs, *J. Am. Chem. Soc.*, 2008, **130**, 17774–17789.
- 29 L. C. M. Ngoka and M. L. Gross, *J. Mass Spectrom.*, 2000, **35**, 265–276.





# Travail expérimental

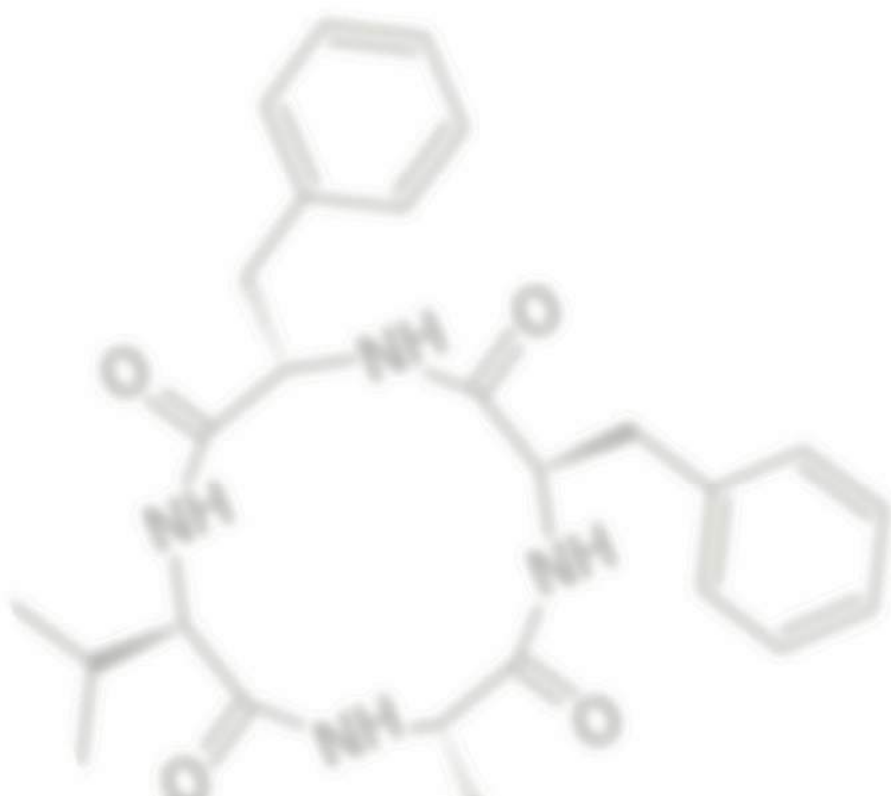
« Engineered: Is it a hold up?  
Doc: It's a science experiment! »  
*Back to the Future*



# 1

*Travail expérimental*

## **Caractérisation du métabolome secondaire de *P. verrucosum***





## 1. ARTICLE 2. CARACTÉRISATION DU MÉTABOLOME SECONDAIRE DE *P. verrucosum*

---

Par l'utilisation de la méthode décrite précédemment, le métabolome secondaire de *P. verrucosum* a pu être caractérisé chimiquement. Ainsi, 98 composés ont été spécifiquement détectés et 18 ont été annotés grâce aux bases de données et à la littérature. Ces annotations ont été vérifiées par la comparaison à des composés de référence où appuyées par l'étude de leurs profils de fragmentation. Ces derniers ont notamment été confrontés aux spectres de fragmentation théoriques générés *in silico* par CFM-ID. L'optimisation et la réalisation d'un réseau moléculaire a permis la mise en évidence de la forte similarité entre les spectres MS/MS d'un composé inconnu et d'un métabolite annoté en tant que fungisporine A hydrolysée. Des études structurales par spectrométrie de masse en tandem ont permis d'émettre une hypothèse quant à la structure de ce nouveaux métabolite secondaire.

Les pages suivantes présentent l'Article 2, publié dans un numéro spécial de *Journal of Chromatography B* consacré aux recherches non ciblées de produits naturels. Outre la caractérisation d'un nouveau métabolite secondaire, cet article présente toutes les informations concernant les fragmentations des composés inconnus par spectrométrie de masse en modes CID et HCD. Le partage de ces informations contribue à surpasser la problématique du manque de données MS/MS quant aux métabolites secondaires fongiques. Ces informations peuvent également être utilisées pour de futures analyses du métabolome secondaire de *Penicillium*.

En deuxième partie de ce chapitre sont présentés les efforts d'optimisation des données de spectrométrie de masse et des paramètres des réseaux moléculaires. En effet, la réalisation des réseaux moléculaires peut s'avérer complexe dans le cadre de l'analyse de jeux de données complexes tels que ceux issus d'extraits fongiques. Cet aspect n'est pas détaillé dans l'article 2 mais fait partie intégrante du travail expérimental réalisé dans cette thèse. Enfin, des éléments non présentés dans l'article quant à l'interprétation des réseaux moléculaires sont détaillés en troisième partie. Ces informations supplémentaires permettent de souligner certaines limites de l'outil de génération des réseaux moléculaires qui seront évoquées dans la discussion de cette thèse.







Contents lists available at ScienceDirect

# Journal of Chromatography B

journal homepage: [www.elsevier.com/locate/chromb](http://www.elsevier.com/locate/chromb)



## Evidencing 98 secondary metabolites of *Penicillium verrucosum* using substrate isotopic labeling and high-resolution mass spectrometry

Thaïs Hautbergue<sup>a,b</sup>, Olivier Puel<sup>a,\*</sup>, Souria Tadriss<sup>a</sup>, Lauriane Meneghetti<sup>a,b</sup>,  
Michel Péan<sup>c,d,e</sup>, Marcel Delaforge<sup>f</sup>, Laurent Debrauwer<sup>a,b</sup>, Isabelle P. Oswald<sup>a</sup>,  
Emilien L. Jamin<sup>a,b</sup>

<sup>a</sup> Toxalim, Université de Toulouse, INRA, INP-ENVT, INP-El-Purpan, Univ. Toulouse 3 Paul Sabatier, 31027 Toulouse, France

<sup>b</sup> Axiom Platform, MetaToul-MetaboHUB, National Infrastructure for Metabolomics and Fluxomics, 31027 Toulouse, France

<sup>c</sup> Groupe de Recherches Appliquées en Phytotechnologie, CEA, IBEB, Cadarache, FR 13108 Saint-Paul-les-Durance, France

<sup>d</sup> UMR Biologie Végétale et Microbiologie Environnementale, CNRS, FR 13108 Saint-Paul-les-Durance, France

<sup>e</sup> Université d'Aix-Marseille, FR 13007 Marseille, France

<sup>f</sup> CNRS, URA 2096, CEA Saclay, FR 91191 Gif sur Yvette, France

### ARTICLE INFO

#### Article history:

Received 27 October 2016

Received in revised form 6 March 2017

Accepted 10 March 2017

Available online xxx

#### Keywords:

Fungus

*Penicillium verrucosum*

Mass spectrometry

Stable isotope labeling

Secondary metabolite

Molecular networking

### ABSTRACT

Industrial applications of fungal compounds, coupled with the emergence of fungal threats to natural ecosystems and public health, have increased interest in filamentous fungi. Among all pathogenic fungi, *Penicillium verrucosum* is one of the most common mold-infecting stored cereals in temperate regions. However, it is estimated that 80% of fungal secondary metabolites remain unknown. To detect new *P. verrucosum* compounds, an untargeted metabolomic approach was applied to fungus grown on wheat grains labeled with stable isotopes: (i) natural grains (99% <sup>12</sup>C); (ii) grains enriched with 97% of <sup>13</sup>C; and (iii) grains enriched with 53% of <sup>13</sup>C and 97% of <sup>15</sup>N. Analyses performed by high-performance liquid chromatography coupled with high-resolution mass spectrometry (HPLC–HRMS) enabled the specific detection of fungal metabolites, and the unambiguous characterization of their chemical formulas. In this way, 98 secondary metabolites were detected and their chemical formulas were determined. Of these, only 18 identifications could be made based on databases, the literature and mass spectrometry fragmentation experiments, with the result that 80 were totally unknown. Molecular networks were generated to analyze these results, leading to the characterization by MS<sup>n</sup> experiments of a new fungisporin produced by *P. verrucosum*. More generally, this article provides precise mass spectrometric data about all these compounds for further studies of the *Penicillium* metabolome.

© 2017 Elsevier B.V. All rights reserved.

### 1. Introduction

Filamentous fungi are well-known for their ability to produce a wide range of secondary metabolites. These molecules are mostly involved in fungal defense and environmental adaptation systems [1]. Since the discovery of penicillin as the first antibiotic of fungal origin used for human health [2], natural compounds have attracted great interest in a variety of industrial areas. Today, 42% of the metabolites adapted for human health applications are derived from fungal compounds [3]. In the agri-food field, they are commonly used as plant-growth hormones [4] or food additives [5]. On the other hand, fungal secondary metabolites known as ‘myco-

toxins’ produce deleterious effects directly on their vegetal hosts, but also indirectly through human food or livestock feed contamination. Mycotoxins are considered the most important chronic dietary risk factor, higher than pesticide residues [6]. Various types of toxicity are associated with fungal metabolites. Some of them display carcinogenic potential, such as aflatoxin, [7] or nephrotoxicity, such as ochratoxin A [8]. Others are estrogenic mycotoxins such as zearalenone, which is involved in reproductive disorders of farm animals and in hyperestrogenic syndromes in humans [9]. Moreover, contaminations by mycotoxins produce major economic consequences, due to the destruction of crops and stocks [10]. Today, although several hundreds of secondary metabolites are characterized as mycotoxins and about 30 are considered of human and animal health concern, only six of them are regulated. Furthermore, fungal genome observations suggest that 80% of their

\* Corresponding author.

E-mail address: [olivier.puel@inra.fr](mailto:olivier.puel@inra.fr) (O. Puel).

secondary metabolome remains unknown, highlighting the great proportion of compounds still to be discovered [11].

*Penicillium verrucosum* is one of the most common molds infecting stored cereals in temperate regions [12]. Its ability to grow at mild temperatures can lead to the infection of grains during their storage, particularly in case the drying or re-moistening process fails [8,13]. In Canada, annual economic losses due to the elimination and destruction of infected cereals stocks are estimated to be over 260 million Canadian dollars [14]. Moreover, ochratoxin A, the major mycotoxin of *P. verrucosum*, is considered to be the most potent renal carcinogen of natural origin [15], and is classified as a possible human carcinogen (Group 2B) [16]. According to some studies, ochratoxin A could be the cause of Balkan Endemic Nephropathy (BEN), a fatal disease first recorded in 1956 [17–19]. Some other secondary metabolites of *P. verrucosum* have been characterized, such as citrinin (nephrotoxic) and ochratoxin B (an unchlorinated intermediate of ochratoxin A biosynthetic pathway), which are considered as mycotoxins, and some metabolites belonging to the families of verrucolones and verrucines [20].

Among the relevant analytical techniques for the characterization of secondary metabolites, untargeted metabolomics using High-Resolution Mass Spectrometry (HRMS) are particularly interesting [21]. This technique enables the detection of both known and unknown metabolites, and the determination of their exact mass, and contributes to their structural identification. However, data interpretation remains complex due to the non-specific detection of compounds including both metabolites and contaminants, and the difficulty of achieving unambiguous chemical formulas, even with highly accurate mass measurement [22]. To overcome these difficulties, a number of original analytical approaches using stable isotope labeling (SIL) have been developed. Bueschl et al. pointed out the benefits provided by this strategy for the specific detection of the metabolites, as well as for the characterization of the global biochemical composition of biological samples [23]. One SIL strategy consists of growing a fungus on the non-labeled glucose medium ( $^{12}\text{C}$ ) as well as on the fully  $^{13}\text{C}$  labeled glucose. The two extracts are then pooled and analyzed by High-Performance Liquid Chromatography (HPLC) coupled with HRMS [24,25]. The detection of two mirror isotopic distributions enables the rapid determination of the number of carbon atoms in all the detected fungal metabolites, facilitating the determination of their chemical formula. However, for high-molecular weight compounds, a single isotope labeling might not be enough to obtain a unique chemical formula and another stable isotope labeling would be required.

Moreover, fungi modulate the synthesis of their secondary metabolites depending on their needs for the medium colonization. This principle, commonly called OSMAC (One Strain Many Compounds), means that every small change in growth conditions can completely alter the fungal secondary metabolism. This implies that the fungal metabolome should be studied in culture conditions as close as possible to its natural growth environment [26]. As stressed by Hallas-Møller et al. [27], the study of cereal-associated species such as *P. verrucosum* has indeed been chemically investigated, but only by outdated analytical systems and on laboratory culture media. Thus, a SIL method based on the growth of plant pathogens on labeled cereals was previously developed in our laboratory [28]. To avoid the formula ambiguity for high-molecular weight metabolites, two isotopic labelings were used by the HPLC-HRMS analysis of carbon 13 ( $^{13}\text{C}$ ) and nitrogen 15 ( $^{15}\text{N}$ ) labeled wheats. Below, following the successful validation of the method on the well-documented metabolism of *Aspergillus fumigatus* [28], we present the application to the saprophytic fungus *P. verrucosum*, with the aim of detecting unknown metabolites produced during wheat grain contamination.

## 2. Materials and methods

### 2.1. Chemicals and reagents

The formic acid, and solvents used for sample preparation (ethyl acetate and methanol) and for the analyses (acetonitrile) were purchased from Fisher Scientific (Thermo Fisher Scientific, Illkirch, France). The ethyl acetate and methanol were HPLC grade and the acetonitrile was MS (mass spectrometry) grade. Ultra-pure water was generated from a Milli-Q system (Millipore, Saint Quentin en Yvelines, France). Standards of ochratoxin A, citrinin and cyclopiazonic acid were purchased from Sigma (Sigma-Aldrich, Saint-Quentin Fallavier, France), and ochratoxin B was purchased from Fermentek (Fermentek Ltd, Jerusalem, Israel).

### 2.2. Identification of *Penicillium verrucosum*

The study was performed on the strain NRRL 5571. Morphological identification was performed as recommended in the monograph of the *Penicillium* genus [29]. Genomic identification was conducted after DNA extraction. After 5 days of culture on YES medium (2.5 g of yeast extract, 10 g of glucose and 500 mL of distilled water), fungal DNA was extracted. After grinding, the mycelium was incubated in an extraction buffer (CTAB 1%, Tris-HCl pH 8 100 mM, EDTA pH8 20 mM and NaCl 1.4 M), Chloroform, phenol, and isopropanol were successively added. The pellet was washed with ethanol 70% and suspended in water. Amplification was performed with ITS (ITS 4 and 5) and  $\beta$ -tub (a and b) primers according to Adjovi et al. [30] and the PCR result was subjected to gel electrophoresis. Each isolated fragment was then purified with a GenElute PCR Clean-Up Kit (Sigma-Aldrich) and the DNA concentrations were measured with Nanodrop (ND-1000 program, Thermo Fisher Scientific, Illkirch, France). Then, they were sequenced by the dye terminator technology on an AB3130XL sequencer (Applied Biosystems). The DNA sequences of ITS and  $\beta$ -tubulin genes were finally compared by sequence alignment and deposited in GenBank under KY347870 and KY347871 accession numbers respectively.

### 2.3. Production of $^{13}\text{C}$ - and $^{13}\text{C}/^{15}\text{N}$ labeled wheats

Uniform isotopic enrichments were performed on the variety Caphorn of *Triticum aestivum* as previously described [31]. Briefly, two different enrichments were achieved in hermetically sealed chambers (750 L) during a complete growth cycle, i.e. 6 months, with  $^{13}\text{C}$ -labeled  $\text{CO}_2$  and  $^{15}\text{N}$ -labeled nitrate and ammonium salts: one nearly 100%  $^{13}\text{C}$  enriched wheat (referred to as  $^{13}\text{C}$  wheat) and one nearly 50%  $^{13}\text{C}$  and nearly 100%  $^{15}\text{N}$  wheat (referred to as  $^{13}\text{C}/^{15}\text{N}$  wheat). The isotopic enrichments were measured with a Delta V Advantage isotope ratio mass spectrometer (Thermo Fisher Scientific, Illkirch, France):  $96.8\% \pm 0.15\%$  for the  $^{13}\text{C}$  wheat,  $53.4\% \pm 0.3\%$   $^{13}\text{C}$  and  $96.8\% \pm 0.07\%$   $^{15}\text{N}$  for the  $^{13}\text{C}/^{15}\text{N}$  wheat.

### 2.4. Fungal strain and growth conditions

The natural and labeled wheat grains were soaked in water and autoclaved twice. The water activity ( $a_w$ ) of the grains was then measured with an HC2-AW measurement device (Rotronik AG, Basserdorf, Switzerland), and sterile water was added where necessary to reach an  $a_w$  of 0.98 before inoculation. Thirty grams of each type of sterilized labeled wheat grains were placed in thin layers into sterile 140 mm diameter Petri dishes (Sarstedt, Marnay, France) to create a total of 4 culture media: one dish with  $^{13}\text{C}$  wheat, one dish with  $^{13}\text{C}/^{15}\text{N}$  wheat and two dishes with  $^{12}\text{C}$  wheat. The first three plates were inoculated with 250  $\mu\text{L}$  of a spore solution (at  $10^5$  spore/mL) prepared from the PDA culture of the fungus. The

second  $^{12}\text{C}$  wheat Petri plate was used as a control and incubated in the same conditions as the other samples but without fungus. The fungal cultures were then placed in the dark without shaking for 1 month at  $25^\circ\text{C}$ .

### 2.5. Metabolite extraction

After incubation, the content of each Petri dish (wheat grains and mycelium) was extracted with 200 mL of ethyl acetate. After 36 h of shaking on a shaking Table IKA Labor technik HS501 (IKA, Staufen, Germany) at 180 rpm, the extract was filtered and evaporated to dryness. The resulting residue was dissolved in 400  $\mu\text{L}$  of methanol and filtered with 0.45  $\mu\text{m}$  Whatman filters (Whatman, GE Healthcare, Kent, U.K.) before HPLC–HRMS analysis.

### 2.6. Analytical parameters

The extracts were analyzed by High-Performance Liquid Chromatography (HPLC) using an Ultimate 3000 system (Thermo Fisher Scientific, Les Ulis, France). A gradient program of 0.1% formic acid in water (phase A) and 100% acetonitrile (phase B) was used as follows (flow rate, 0.2 mL/min): 0 min 20% of B, 30 min 50% of B, from 35 to 45 min 90% of B, from 50 to 60 20% of B. A volume of 10  $\mu\text{L}$  was injected on a reversed-phase Luna C18 column (150 mm  $\times$  2.0 mm, 5  $\mu\text{m}$ ) (Phenomenex, Torrance, CA, U.S.A.). HPLC was coupled to a LTQ Orbitrap XL hybrid high-resolution mass spectrometer (Thermo Fisher Scientific, Les Ulis, France) using electrospray ionization (ESI) in the positive and negative modes. The ionization parameters were as follows: spray voltage +4.5 kV, capillary temperature  $350^\circ\text{C}$ , sheath gas ( $\text{N}_2$ ) flow rate 40 au (arbitrary units), auxiliary gas ( $\text{N}_2$ ) flow rate 6 au for the positive mode, and spray voltage  $-3.7$  kV, capillary temperature  $350^\circ\text{C}$ , sheath gas flow rate 30 au ( $\text{N}_2$ ), auxiliary gas flow rate 10 au ( $\text{N}_2$ ) for the negative mode. The full HRMS spectra were acquired at a resolution of 60,000 from  $m/z$  50 to  $m/z$  800. The high-resolution mass analyzer was calibrated in each ionization mode based on the supplier's protocol and calibration mixtures (Thermo Fisher Scientific, Les Ulis, France). The targeted MS/MS spectra were performed according to the following parameters: [Parameters for MS/MS by collision-induced dissociation (CID): resolution = 7500,  $\text{MS}^n$  maximum ion time = 200 ms, isolation width = 1.5 Da, Act. Q = 0.250 and activation time = 30 ms]. [Parameters for MS/MS by higher energy collision dissociation (HCD): resolution = 7500,  $\text{MS}^n$  maximum ion time = 200 ms, isolation width = 1.5 Da and activation time = 30 ms]. HCD and CID energies were adjusted to obtain appropriate fragmentation and maximum structural information. Since the analyses were performed to identify metabolites from a strictly qualitative point of view, neither quality control samples, nor replicates were necessary.

Analysis by HPLC coupled with diode-array-detection (Ultimate 3000 system, Thermo Fisher Scientific, Les Ulis, France) were performed to identify anacine. The chromatographic system and parameters were the same as for HPLC–HRMS analyses. The presence of anacine was monitored at a wavelength of 220 nm.

### 2.7. Identification of metabolites

Comparisons of proposed  $^{12}\text{C}$ ,  $^{13}\text{C}$  and  $^{13}\text{C}^{15}\text{N}$  formulas were performed thanks to the MassCompare program developed in-house (<http://www.toulouse.inra.fr/axiommm>) using a measurement accuracy of 5 ppm. C, H, O, N, S, P, Cl, K and Na were the only accepted elements.

Identifications of metabolites were classified according to the Metabolomics Standard Initiative [32]. Metabolites identified at level 1 in our results displayed the same retention time, chemical

formula and MS/MS fragmentation pattern as the standard compound (Section 2.1) measured in a parallel analysis in the same analytical conditions. The metabolites identified at level 2 in our results displayed an MS/MS or UV spectrum in accordance with the bibliography. The metabolites annotated as compounds known to be produced by *P. verrucosum* and presenting coherent MS/MS spectra were identified at the level 3A. The metabolites annotated as compounds never characterized on *P. verrucosum* and presenting coherent MS/MS spectra were identified at the level 3B. Complementary *in silico* fragmentations were performed using CFM-ID [33] and compared with experimental MS/MS spectra of annotated metabolites of level 3.

A molecular network generated using GNPS [34] was created from MS/MS spectra in the CID mode of all annotated secondary metabolites. The software parameters were set as follows: ion mass tolerance = 0.0036 Da, cosine score = 0.7, minimum match fragment ion = 1, maximum connected components = 100 and Top-K = 100, filter peaks removed. The resulting molecular network was visualized in Cytoscape 2.8.2 [35].

## 3. Results and discussion

### 3.1. Certification of *Penicillium verrucosum*

The identification of fungal strains could be problematic for secondary metabolome studies due to the lack of robust classification criteria [20], despite the availability of genetic classification. For example, species such as *Penicillium nordicum* and *P. verrucosum* were considered as *Penicillium viridicatum* for some time and some of their secondary metabolites are still referenced as *P. viridicatum* compounds, therefore misleading studies on these fungi. Given the confusion between these similar fungi, it was crucial for our study to ensure the identification of *P. verrucosum*. Three different approaches were investigated. Firstly, a morphological characterization was achieved and displayed both macroscopic and microscopic *Penicillium* properties, namely terverticillate conidiophores, rough-walled and globose conidia. At this stage, no differentiation could be made between *P. verrucosum* and *P. nordicum* as these two species display the same characteristics. Then, cultures on YES agar were performed and showed a red-brown reverse color characteristic of *P. verrucosum* under some sporulation areas. This observation is the only morphological point enabling differentiation between *P. verrucosum* and *P. nordicum* [29]. Secondly, the fungal secondary metabolite analysis achieved by HPLC–HRMS led to the identification of ochratoxin A and citrinin, which were validated by comparison with the standards. Ochratoxin A is known to be produced by both *P. nordicum* and *P. verrucosum* but, according to the *Penicillium* classification of Larsen et al. [20], only *P. verrucosum* is known to produce citrinin in addition to ochratoxin A. Finally, a genetic identification was conducted. Since Samson et al. [36] estimated that  $\beta$ -tubulin sequences are relevant species markers, we compared the  $\beta$ -tubulin DNA sequences of the studied fungus with the standard sequences of *P. verrucosum* and *P. nordicum*. As a result,  $\beta$ -tubulin sequences showed strong similarities with those of *P. verrucosum*. ITS DNA sequences are well-conserved sequences also used as species markers in fungal studies. Comparison of the ITS sequences confirmed the strong similarity between the studied strain (NRRL 5571) and the standard sequences of *P. verrucosum*. To conclude, the three identification steps (morphological observations, metabolite characterizations and genetic validation) were in accordance with the identification of *P. verrucosum*. NRRL 5571 was also classified as *P. verrucosum* by Larsen et al. in a study characterizing 48 *P. verrucosum* strains [20].

### 3.2. HPLC–HRMS analysis of the three wheat cultures

The characterization of the fungal secondary metabolome was carried out according to the protocol previously developed in our group [28]. Natural wheat,  $^{13}\text{C}$  wheat and  $^{13}\text{C}^{15}\text{N}$  wheat were infected with *P. verrucosum* under adapted culture conditions (*i.e.* culture in the dark at 25 °C for 30 days). A control sample, corresponding to non-infected natural wheat, was placed in the same environment. The secondary metabolites were then extracted and the samples analyzed separately by HPLC–HRMS. The extracts were not pooled in order to limit the complexity of MS spectra due to the complex isotopic compositions used in our protocol and to avoid possible overlapping signals despite the high-resolution power provided by the Orbitrap analyzer. The total ion current chromatograms obtained in the positive and negative ionization modes are presented in Fig. 1. First, the chromatogram and mass spectrum obtained from the  $^{13}\text{C}^{15}\text{N}$  labeling experiment were analyzed to specifically detect all the secondary metabolites. The isotopic labeling with 50%  $^{13}\text{C}$  and 100%  $^{15}\text{N}$  provides an easily recognizable signal, since it generates a specific ion pattern corresponding to a normal distribution of  $^{12}\text{C}/^{13}\text{C}$  on the mass spectrum (Fig. 2C). These specific signals could be exclusively attributed to either the labeled substrate or the fungal metabolites, enabling the removal of any contaminant or chemical used during incubation or extraction. Here, 91 and 130 specific isotopic patterns were detected by positive (ESI+) and negative (ESI–) electrospray ionization modes respectively. Among these signals, 33 adducts and in-source fragment ions were rejected in the ESI+ mode (68 were rejected for the ESI– mode) in order to consider only molecular ions for each secondary metabolite [37]. This led to a final listing of 58 signals in the positive ionization mode and 62 in the negative mode. The same approach was applied to the  $^{13}\text{C}$  labeling experiment, in which secondary metabolites displayed another specific pattern (Fig. 2B) due to the 3% of remaining  $^{12}\text{C}$ . For each previously detected  $^{13}\text{C}^{15}\text{N}$  labeled ions, the  $^{13}\text{C}$  equivalent signal was found in HRMS spectra. Each secondary metabolite was finally searched in the HRMS spectra obtained from the natural  $^{12}\text{C}$  experiment (Fig. 2A). To validate the detection of similar metabolites in the three parallel HPLC–HRMS analyses ( $^{12}\text{C}$ ,  $^{13}\text{C}$  and  $^{13}\text{C}^{15}\text{N}$ ), chromatographic retention times and MS/MS spectra were compared. As displayed in Fig. 2, MS/MS spectra were conducted on precursor ions detected at the same retention time (26.7 min in the example given in Fig. 2), evidencing that the same metabolite was detected in the three labeled substrates. Finally, wheat molecules were rejected by comparing extract and control samples. In this way, 3 and 13 compounds were eliminated from the ESI+ and ESI– results, respectively. A final list of the 98 secondary metabolites detected from *P. verrucosum* is reported in Table 1. The metabolites detected both in ESI+ and ESI– are displayed as a single entry in Table 1.

For further identification of the detected metabolites, the first step was the determination of their chemical formulas. Each secondary metabolite, measured  $m/z$  in  $^{12}\text{C}$ ,  $^{13}\text{C}$  and  $^{13}\text{C}^{15}\text{N}$  extracts was compared using MassCompare, a software developed in-house. This tool calculates all the possible chemical formulas for each  $^{12}\text{C}$ ,  $^{13}\text{C}$  and  $^{13}\text{C}^{15}\text{N}$  metabolites, selects the common ones, and calculates the number of carbon and nitrogen atoms by subtractions between  $^{12}\text{C}$ ,  $^{13}\text{C}$  and  $^{13}\text{C}^{15}\text{N}$   $m/z$  values. In this way, MassCompare provides the unique chemical formula of each metabolite. Concerning the example illustrated in Fig. 2, the mass difference between the  $m/z$  ratio of the  $^{12}\text{C}$  ion ( $m/z$  370.1285) and the  $^{13}\text{C}$ -labeled ion ( $m/z$  390.1951) was used to determine the number of carbon atoms ( $n_{\text{C}} = 20$ ). The mass differences between the  $m/z$  of the  $^{12}\text{C}$  ion, the  $^{13}\text{C}$ -labeled ion and the  $^{13}\text{C}/^{15}\text{N}$ -labeled ion ( $m/z$  382.1667) enabled the calculation of the number of nitrogen atoms ( $n_{\text{N}} = 1$ ) by MassCompare. With this information, the software determined a single chemical formula per metabolite ( $\text{C}_{20}\text{H}_{19}\text{O}_6\text{N}$  for the exam-

ple of Fig. 2). As displayed in Table 1, the metabolite formulas are almost homogenous in terms of heteroatom composition with a majority of oxygenated structures, and only seven non-oxygenated metabolites. Over half of the 98 metabolites contain at least one nitrogen atom. These observations were in agreement with the general composition of fungal secondary metabolites. These extrolites are synthesized by two major types of enzymes. The first, polyketide synthases (PKS), synthesize secondary metabolites from acetyl or malonyl groups to generate oxygenated compounds [38]. The second, nonribosomal peptide-synthetases (NRPS) ensure the incorporation of amino acids, leading to the synthesis of nitrogenous secondary metabolites. Among all detected metabolites, only three display other heteroatoms: one chlorinated and two sulfured formulas. The chlorinated compound was identified as the well-known ochratoxin A. To the best of our knowledge, it is the first time that sulfur containing metabolites has been detected in *P. verrucosum*.

### 3.3. Dereplication of known secondary metabolites

Experimental MS/MS spectra were compared with the open-access GNPS library of natural products [34]. In this study, only one match was found but the MS/MS spectra were not sufficiently similar to allow its identification. Alternatively, the database of natural compounds Antibase (2012) [39], coupled with the literature on fungal secondary metabolites, was used to successfully annotate several compounds according to their chemical formulas.

Ochratoxin A (compound 1 in Fig. 3) is considered as the major toxin of *P. verrucosum* involved in the sanitary and economic troubles caused by this fungus [40,41]. This metabolite was suspected to be detected at 32.4 min at  $m/z$  404.0897 with a chlorinated isotopic pattern after positive ionization of the  $^{12}\text{C}$  sample. Separate observation of the  $^{12}\text{C}$  spectrum, with consideration of the chlorine atom, led to eight possibilities of chemical formulas using the nitrogen rules. The comparison of the three lists of possible formulas from the three different wheat cultures and calculation of the number of carbon and nitrogen atoms finally led to a unique chemical formula:  $\text{C}_{20}\text{H}_{19}\text{O}_6\text{NCl}$ . This result was confirmed by corresponding data obtained in the negative mode. The comparison of chromatographic and MS/MS patterns of the metabolite with those of the standard of ochratoxin A validated the identification at level 1 [32] (see Section 2.7). Similarly, analysis of the standard of citrinin, a well-known *P. verrucosum* metabolite (compound 3 in Fig. 3) [20,42], confirmed its production by the studied strain. Ochratoxin B (compound 2 in Fig. 3) was also identified at a level 1 ( $\text{C}_{20}\text{H}_{20}\text{O}_6\text{N}$ ) at  $R_{\text{T}} = 26.7$  min, in accord with other works which noticed the excretion of this dechlorinated intermediate of the biosynthesis pathway of ochratoxin A by *P. verrucosum* [20,29]. Cyclopiiazonic acid, never observed in *P. verrucosum*, was proposed for annotation of the metabolite  $\text{C}_{20}\text{H}_{20}\text{O}_3\text{N}_2$  but the chromatographic and MS/MS results obtained from the reference compound ruled out this hypothesis.

Since many natural products are unavailable as reference compounds [37], some secondary metabolites already known to be produced by *P. verrucosum* were detected without it being possible to validate their identifications. The quinazolines verrucine A (compound 6 in Fig. 3) [20,43], verrucine B (compound 7 in Fig. 3) [20,43] and verrucine F (compound 9 in Fig. 3) [44] were annotated. Their MS/MS spectra displayed a loss of ammonia and a consecutive loss of carbon monoxide in accordance with the structures of verrucines A, B and F. Furthermore, the fragmentation patterns of verrucines A and B have already been reported [43], allowing the identification of verrucines A and B at a level 2. Since standards were not available, *in silico* fragmentations were also performed using CFM-ID [33]. Although it should not be considered as proof of identification, this innovative tool provides relevant supporting

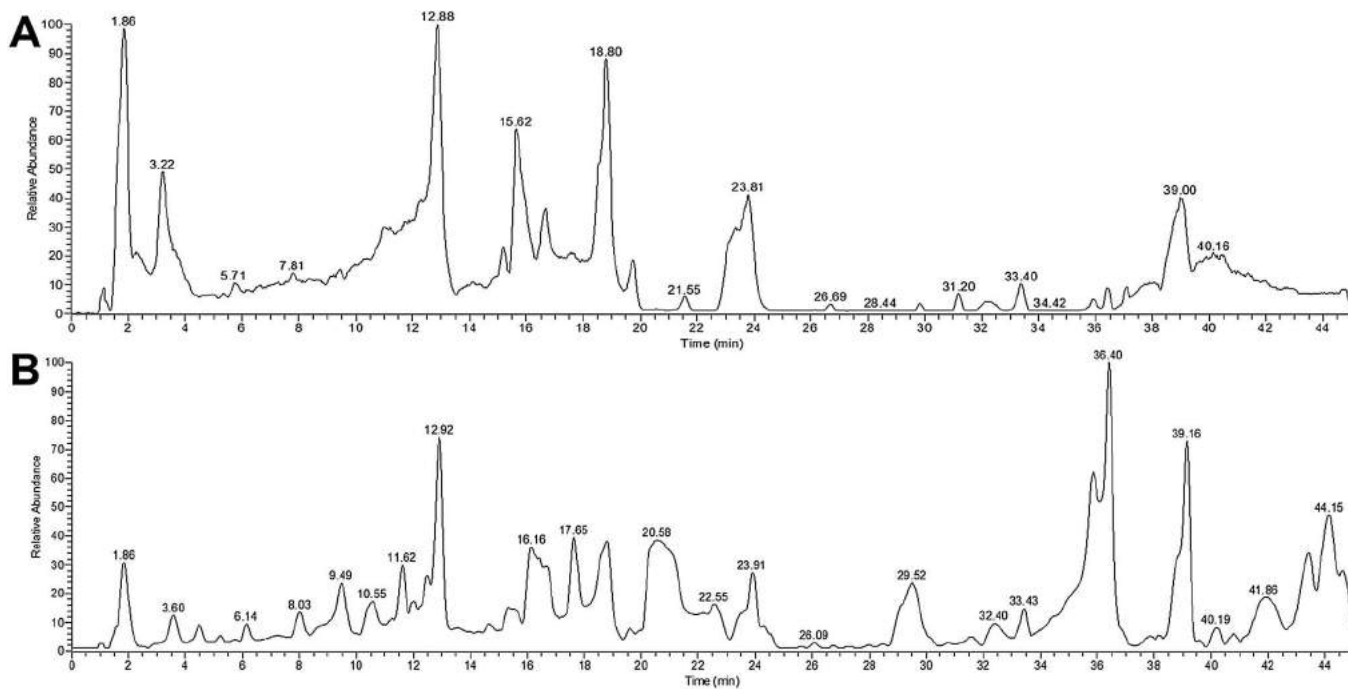


Fig. 1. Total ion current chromatograms obtained in the positive (A) and negative (B) ionization modes from the extract of wheat grains infected by *Penicillium verrucosum*.

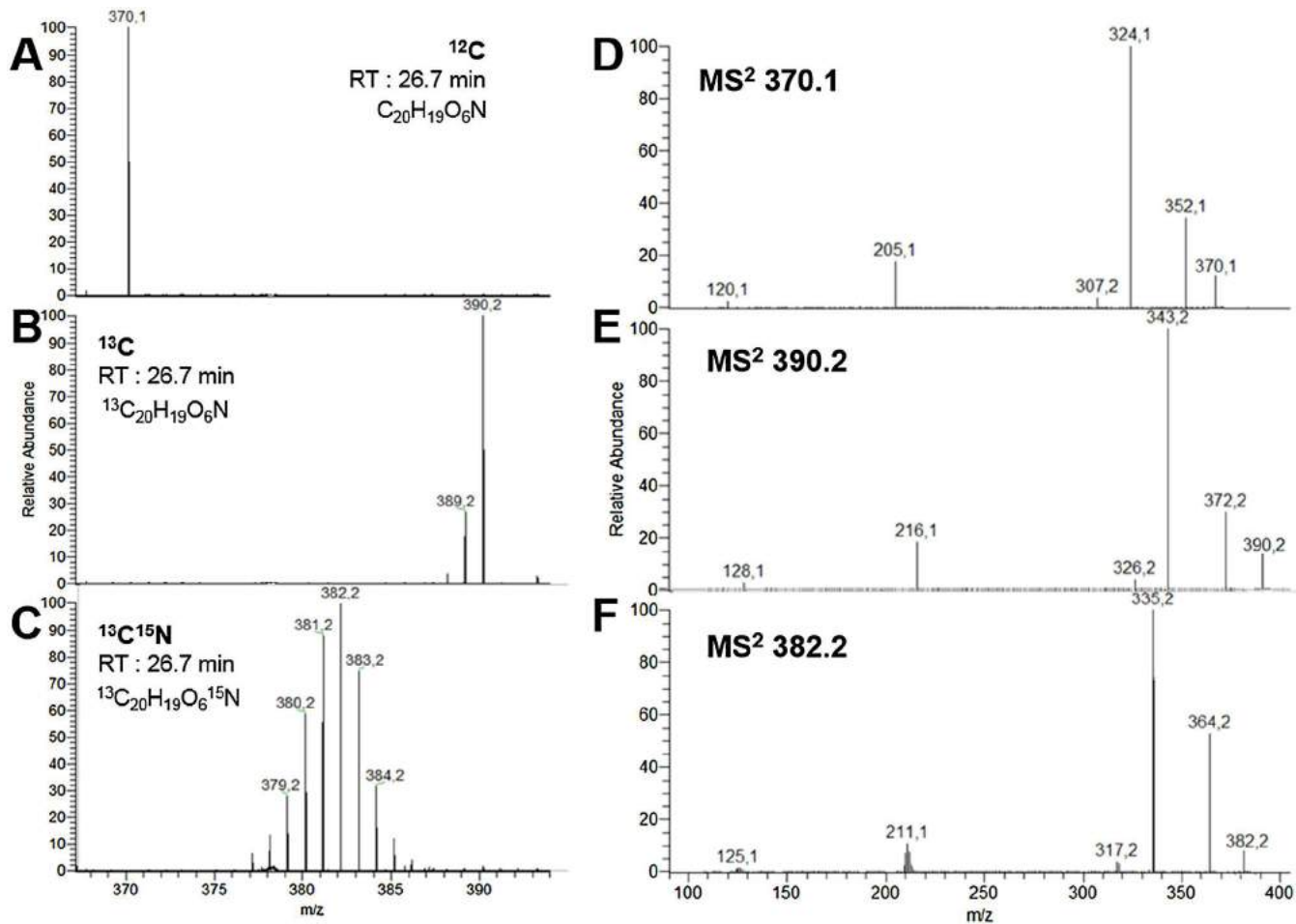


Fig. 2. Detection of ochratoxin B: mass spectra detected at  $R_T = 26.7$  min after positive ESI-HPLC-MS from extract of *Penicillium verrucosum* grown (A) on natural 99%  $^{12}\text{C}$  wheat grains, (B) on grains labeled with 97% of  $^{13}\text{C}$  and (C) on grains labeled with 53% of  $^{13}\text{C}$  and 97% of  $^{15}\text{N}$ . Right-hand panels display the corresponding CID MS/MS spectra from culture (D) on natural 99%  $^{12}\text{C}$  wheat grains, (E) on grains labeled with 97% of  $^{13}\text{C}$  and (F) on grains labeled with 53% of  $^{13}\text{C}$  and 97% of  $^{15}\text{N}$ .

**Table 1**  
List of the secondary metabolites detected in *Penicillium verrucosum* grown on wheat grains.

molecular formula <sup>a</sup>	Identifier <sup>f</sup>	proposed identification	molecular formula <sup>a</sup>	Identifier <sup>f</sup>	proposed identification	molecular formula <sup>a</sup>	Identifier <sup>f</sup>	proposed identification
C <sub>6</sub> H <sub>8</sub> O <sub>3</sub>	PVer_129.05.7.6	Verrucolone <sup>d</sup>	C <sub>18</sub> H <sub>24</sub> O <sub>7</sub>	PVer_351.14.29.1	Ochratoxin A <sup>b</sup>	C <sub>24</sub> H <sub>28</sub> N <sub>4</sub>	PVer_373.24.16.0	Fungisporin C <sup>e</sup>
C <sub>6</sub> H <sub>10</sub> O <sub>4</sub>	PVer_147.06.3.1		C <sub>18</sub> H <sub>24</sub> O <sub>7</sub>	PVer_351.14.24.1		C <sub>25</sub> H <sub>30</sub> N <sub>4</sub>	PVer_387.25.18.0	
C <sub>11</sub> H <sub>18</sub> O <sub>5</sub>	PVer_231.12.5.6	LL-P880β <sup>c</sup>	C <sub>18</sub> H <sub>28</sub> O <sub>2</sub>	PVer_277.21.39.4	Ochratoxin B <sup>b</sup>	C <sub>25</sub> H <sub>30</sub> N <sub>4</sub>	PVer_385.24.35.7	Hydrolyzed fungisporin A <sup>c</sup>
C <sub>12</sub> H <sub>12</sub> O <sub>5</sub>	PVer_235.06.24.2	Citrinin <sup>b</sup>	C <sub>18</sub> H <sub>32</sub> O <sub>4</sub>	PVer_335.22.36.2		C <sub>25</sub> H <sub>30</sub> N <sub>4</sub>	PVer_385.24.37.2	
C <sub>13</sub> H <sub>13</sub> NO <sub>4</sub>	PVer_248.09.11.4		C <sub>18</sub> H <sub>32</sub> O <sub>4</sub>	PVer_335.22.36.8	C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	PVer_513.21.20.6		
C <sub>13</sub> H <sub>14</sub> O <sub>4</sub>	PVer_235.10.8.3		C <sub>19</sub> H <sub>18</sub> N <sub>2</sub> O <sub>2</sub>	PVer_307.14.38.5	C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	PVer_513.21.24.6		
C <sub>13</sub> H <sub>14</sub> O <sub>4</sub>	PVer_235.10.19.8		C <sub>19</sub> H <sub>18</sub> N <sub>2</sub> O <sub>3</sub> S	PVer_355.11.28.1	C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	PVer_513.21.26.3		
C <sub>13</sub> H <sub>14</sub> O <sub>5</sub>	PVer_251.09.23.5		C <sub>19</sub> H <sub>26</sub> O <sub>6</sub>	PVer_371.15.22.5	C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	PVer_513.21.26.7		
C <sub>13</sub> H <sub>16</sub> O <sub>5</sub>	PVer_251.09.40.0		C <sub>19</sub> H <sub>37</sub> NO <sub>2</sub>	PVer_312.29.31.9	C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	PVer_513.21.27.1		
C <sub>13</sub> H <sub>16</sub> N <sub>2</sub> O <sub>6</sub>	PVer_295.09.4.6		C <sub>20</sub> H <sub>18</sub> NO <sub>6</sub> Cl	PVer_404.09.33.1	C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	PVer_513.21.30.1		
C <sub>13</sub> H <sub>16</sub> O <sub>6</sub>	PVer_267.09.9.2		C <sub>20</sub> H <sub>19</sub> NO <sub>6</sub>	PVer_402.07.33.1	C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	PVer_513.21.31.4		
C <sub>13</sub> H <sub>16</sub> O <sub>6</sub>	PVer_267.09.6.5			PVer_370.13.26.5	PVer_370.13.26.5	C <sub>28</sub> H <sub>36</sub> N <sub>4</sub> O <sub>6</sub>	PVer_523.25.25.8	
C <sub>13</sub> H <sub>16</sub> O <sub>6</sub>	PVer_267.09.21.9		C <sub>20</sub> H <sub>20</sub> N <sub>2</sub> O <sub>3</sub>	PVer_368.11.26.5	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	PVer_511.29.18.0		
C <sub>13</sub> H <sub>20</sub> O <sub>5</sub>	PVer_257.14.13.6			PVer_337.15.38.1	PVer_337.15.38.1	PVer_509.28.18.0	New tetrapeptide	
C <sub>14</sub> H <sub>12</sub> N <sub>4</sub> O <sub>4</sub>	PVer_299.08.3.5		C <sub>20</sub> H <sub>20</sub> N <sub>2</sub> O <sub>4</sub>	PVer_353.15.32.0	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	PVer_511.29.18.7		
C <sub>15</sub> H <sub>11</sub> O <sub>3</sub>	PVer_238.06.11.5	C <sub>20</sub> H <sub>28</sub> O <sub>8</sub>	PVer_395.17.15.0	PVer_509.28.18.7		Hydrolyzed fungisporin B <sup>e</sup>		
C <sub>15</sub> H <sub>19</sub> N <sub>3</sub>	PVer_242.16.16.1	C <sub>20</sub> H <sub>30</sub> O <sub>5</sub>	PVer_349.20.22.5	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>6</sub>	PVer_525.27.13.0			
C <sub>16</sub> H <sub>12</sub> O <sub>3</sub>	PVer_251.07.7.9	C <sub>21</sub> H <sub>18</sub> N <sub>4</sub> O <sub>3</sub>	PVer_375.14.18.7		C <sub>30</sub> H <sub>38</sub> N <sub>4</sub> O <sub>2</sub>	PVer_509.29.26.2		
C <sub>16</sub> H <sub>12</sub> O <sub>3</sub>	PVer_251.07.11.6	C <sub>21</sub> H <sub>18</sub> N <sub>4</sub> O <sub>3</sub>	PVer_373.13.17.0	C <sub>30</sub> H <sub>38</sub> N <sub>4</sub> O <sub>2</sub>	PVer_509.29.25.2			
C <sub>16</sub> H <sub>17</sub> N <sub>3</sub> O <sub>2</sub>	PVer_284.14.22.8	C <sub>21</sub> H <sub>18</sub> N <sub>4</sub> O <sub>4</sub>	PVer_391.14.13.5	C <sub>30</sub> H <sub>38</sub> N <sub>4</sub> O <sub>2</sub>	PVer_509.29.23.0			
C <sub>17</sub> H <sub>14</sub> O <sub>5</sub>	PVer_299.09.11.5	C <sub>21</sub> H <sub>18</sub> N <sub>4</sub> O <sub>4</sub>	PVer_389.12.13.5	C <sub>30</sub> H <sub>38</sub> N <sub>4</sub> O <sub>2</sub>	PVer_509.29.24.4			
C <sub>17</sub> H <sub>34</sub> O <sub>5</sub>	PVer_297.08.11.5		C <sub>21</sub> H <sub>18</sub> N <sub>4</sub> O <sub>4</sub>	PVer_391.14.21.0	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	PVer_548.29.17.4		
	PVer_317.23.41.0	C <sub>21</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	PVer_377.16.18.4	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	PVer_548.29.18.0			
	PVer_289.10.24.4	C <sub>21</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	PVer_377.16.12.8	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	PVer_548.29.19.1			
	PVer_304.11.36.5	C <sub>21</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	PVer_393.15.9.4	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	PVer_564.28.12.6			
	PVer_304.11.38.1	C <sub>21</sub> H <sub>20</sub> N <sub>4</sub> O <sub>4</sub>	PVer_391.14.9.4	PVer_391.14.9.4	PVer_567.09.26.3			
	PVer_304.11.36.4		C <sub>21</sub> H <sub>20</sub> N <sub>4</sub> O <sub>4</sub>	PVer_393.15.15.5	C <sub>31</sub> H <sub>20</sub> O <sub>11</sub>	PVer_567.09.25.1		
	PVer_387.06.3.8	C <sub>21</sub> H <sub>21</sub> N <sub>3</sub> O <sub>3</sub>	PVer_364.16.3.6	C <sub>31</sub> H <sub>40</sub> N <sub>4</sub> O <sub>2</sub>	PVer_523.30.27.7			
	PVer_341.16.20.1	C <sub>21</sub> H <sub>21</sub> N <sub>3</sub> O <sub>3</sub>	PVer_364.16.8.8	C <sub>31</sub> H <sub>40</sub> N <sub>4</sub> O <sub>2</sub>	PVer_523.30.27.4			
	PVer_341.16.15.1	C <sub>21</sub> H <sub>30</sub> N <sub>4</sub> O <sub>4</sub>	PVer_401.22.25.8	C <sub>32</sub> H <sub>36</sub> N <sub>4</sub> O <sub>2</sub>	PVer_507.28.38.5			
	PVer_343.17.12.7	C <sub>21</sub> H <sub>30</sub> N <sub>4</sub> O <sub>4</sub>	PVer_401.22.21.9	C <sub>32</sub> H <sub>42</sub> O <sub>7</sub>	PVer_537.29.39.5			
	PVer_343.17.17.9	C <sub>22</sub> H <sub>27</sub> N <sub>3</sub>	PVer_334.23.15.8	C <sub>32</sub> H <sub>42</sub> O <sub>8</sub>	PVer_553.28.38.7			
	PVer_351.14.18.1	C <sub>22</sub> H <sub>27</sub> N <sub>3</sub>	PVer_332.21.35.6	C <sub>32</sub> H <sub>42</sub> O <sub>8</sub>	PVer_553.28.38.3			
PVer_351.14.29.0	C <sub>23</sub> H <sub>21</sub> N <sub>5</sub> O <sub>3</sub>	PVer_416.17.14.6	C <sub>37</sub> H <sub>42</sub> N <sub>4</sub> O <sub>5</sub>	PVer_623.32.36.3				
PVer_351.14.24.6	C <sub>23</sub> H <sub>23</sub> N <sub>3</sub> O <sub>4</sub>	PVer_406.17.8.8	C <sub>38</sub> H <sub>42</sub> N <sub>4</sub> O <sub>10</sub>	PVer_621.30.36.3				
				PVer_713.28.33.4				

<sup>a</sup> The chemical formulas were calculated after comparison of the *m/z* ratios obtained from the culture of *P. verrucosum* in <sup>12</sup>C wheat, 97% <sup>13</sup>C wheat and 53% <sup>13</sup>C/96% <sup>15</sup>N wheat.

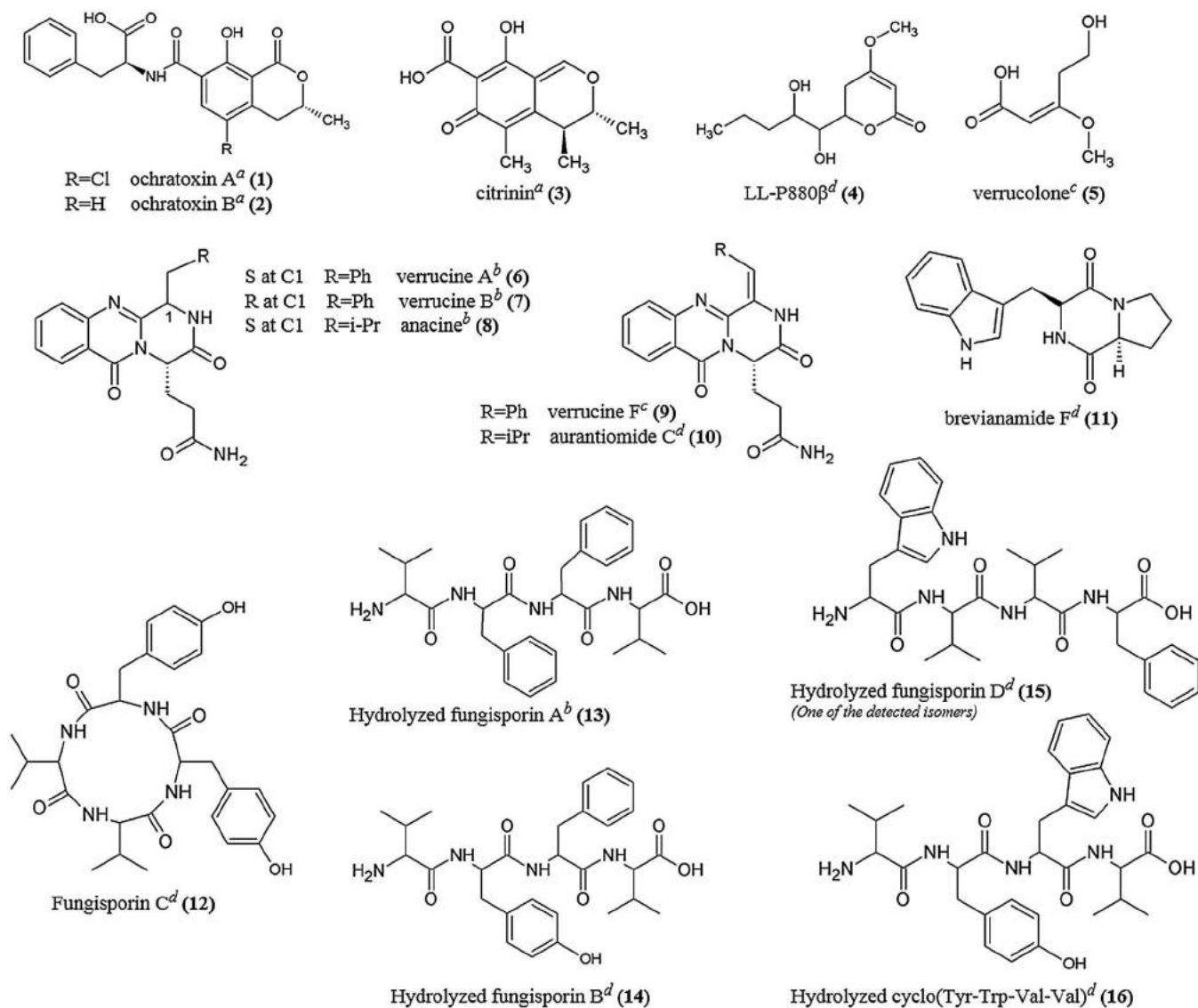
<sup>b</sup> Identification at level 1.

<sup>c</sup> Identification at level 2.

<sup>d</sup> Identification at level 3 of compounds known to be produced by *P. verrucosum*.

<sup>e</sup> Identification at level 3 of compounds never characterized in *P. verrucosum* but in other *Penicillium* species.

<sup>f</sup> Secondary metabolites identifier (PVer.*m/z*.R<sub>T</sub>).

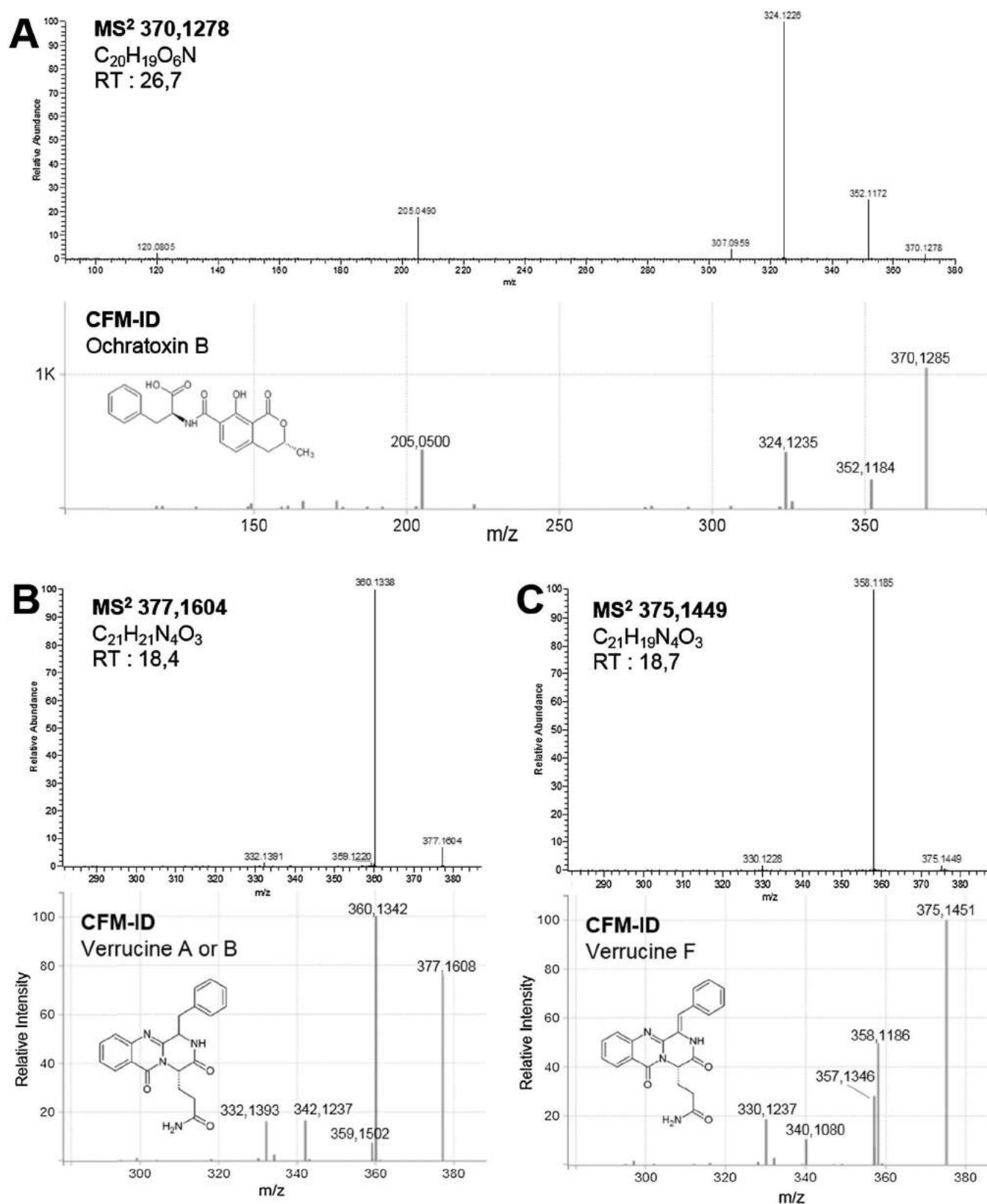


**Fig. 3.** Secondary metabolites identified from *Penicillium verrucosum*. The chemical formulas were calculated after comparison of the  $m/z$  ratios obtained from cultures of *P. verrucosum* in  $^{12}\text{C}$  wheat, 97%  $^{13}\text{C}$  wheat and 53%  $^{13}\text{C}/97\%$   $^{15}\text{N}$  wheat. <sup>a</sup> Identification at level 1, <sup>b</sup> at level 2, <sup>c</sup> at level 3 of compounds known to be produced by *P. verrucosum*, <sup>d</sup> at level 3 of compounds never characterized in *P. verrucosum* but in other *Penicillium* species.

information. The comparison between experimental and simulated MS/MS spectra is shown in Fig. 4A for the formally identified ochratoxin B as an example of successful simulation. Application of CFM-ID to the putative quinazoline structures showed that *in silico* spectra were in accord with the annotation of the three verrucines. Indeed, the same two major fragment ions at  $m/z$  360.2 and  $m/z$  332.2 were detected in both experimental and *in silico* MS/MS spectra of verrucines A and B (Fig. 4B). The same result was obtained with the annotated verrucine F ( $m/z$  375.1450), which shared two fragment ions at  $m/z$  358.2 and  $m/z$  330.2 with the simulated MS/MS spectrum (Fig. 4C). Verrucolone, also known as arabeoic acid (compound 5 in Fig. 3) and referenced as a typical *P. verrucosum* metabolite [45], was attributed to the detected  $\text{C}_6\text{H}_{10}\text{O}_4$  formula. The experimental MS/MS spectrum was compared to CFM-ID simulations and some common fragment ions were observed at  $m/z$  129.0, 117.0, 115.0 and 103.0. Verrucine F and verrucolone were therefore identified with a level 3A (see Section 2.7).

Some secondary metabolites detected in our extracts have already been characterized in *Penicillium* fungi but not in the *P. verrucosum* strain. Unfortunately, no standards of these supposed secondary metabolites were available. In this case, the only supporting information came from the MS/MS or UV spectra available from the literature and from comparison with *in silico* fragmentations. Anacine (compound 8 in Fig. 3) is known to be produced by *Penicillium nordicum*, one of the closest fungi of *P. verrucosum* [20]. This compound is a quinazoline belonging to the verrucines family [46]. In this study, anacine was identified at a level 2 since its UV spectrum was in accordance with the spectrum from literature [43]. Moreover, the unsaturated analog of anacine, namely aurantiomide C (compound 10 in Fig. 3) was proposed for the annotation of  $\text{C}_{18}\text{H}_{21}\text{O}_3\text{N}_4$ . To our knowledge, this secondary metabolite was characterized in *Penicillium aurantiogriseum* [47] but never in *P. verrucosum*. Although its MS/MS spectrum was not available in literature, an *in silico* simulation was performed and compared with the experimental MS/MS fragmentation. Both spectra displayed a major fragment ion at  $m/z$  324.1. In the same way, brevianamide F





**Fig. 4.** Comparison between experimental MS/MS spectra (ESI<sup>+</sup> high-resolution CID) of three secondary metabolites of *Penicillium verrucosum* and the simulated MS/MS spectra of their respective structural hypotheses. (A) Successful *in silico* simulation: experimental MS/MS spectrum of ochratoxin B (upper part) (identified by homology to an authentic standard) and theoretical MS/MS generated by CFM-ID (lower part). (B) MS/MS spectrum obtained from the secondary metabolite  $C_{21}H_{21}N_4O_3$  detected at  $R_T = 18,4$  min (upper part) and *in silico* fragmentation of verrucines A and B using CFM-ID (lower part). (C) MS/MS spectrum obtained from the secondary metabolite  $C_{21}H_{19}N_4O_3$  detected at  $R_T = 18,7$  min (upper part) and *in silico* fragmentation of verrucine F using CFM-ID (lower part).

(compound 11 in Fig. 3), a metabolite of *Penicillium brevicompactum* [39] and a precursor of fumitremorgins in *Aspergillus fumigatus* [48], was proposed for the annotation of  $C_{16}H_{17}O_2N_3$ . In this case again, no MS/MS information was available in the literature, and the *in silico* MS/MS simulation displayed almost all the fragment ions observed on the experimental spectrum. LL-P880 $\beta$  (compound 4 in Fig. 3), another *Penicillium* metabolite whose fragmentation pattern was not published, was suspected based on database researches [39]. This pestalotin analog, also named hydroxy-pestalotin, was reported to be synthesized by *Penicillium* sp. without any precision about the species [39]. Moreover, in a study of the pestalotins family, Rahbaek et al. reported that PC-2, LL-P880 $\gamma$  and two verrucosapyrones displaying very similar structures, were secreted by several strains of *P. verrucosum* [45]. Consequently, *P. verrucosum* likely forms part of the *Penicillium* species able to produce LL-P880 $\beta$ . Finally, several cyclic tetrapeptides of the fungisporin family were suspected to be produced by *P. verrucosum*. Indeed, the metabolite  $C_{28}H_{36}O_6N_4$  showed a fragmentation pattern in accordance with a tetrapeptide structure and was identified as fungisporin C (compound 12 in Fig. 3). This natural compound was previously described by Klitgaard et al. in *Aspergillus nidulans* but the study did not report the fragmentation spectrum [49]. Furthermore, the degradation products of fungisporin A (compound 13 in Fig. 3), fungisporin B (compound 14 in Fig. 3), fungisporin D (compound 15 in Fig. 3) and cyclo-(Tyr-Trp-Val-Val) (compound 16 in Fig. 3) were observed. The MS/MS spectrum of the degradation products of fungisporin A was presented by Ali et al. in their study characterizing fungisporins in *Penicillium chrysogenum* [50]. It was therefore possible to identify at a level 2 the compound 13 (Table 1) as a hydrolysis product of fungisporin A.

### 3.4. Unknown secondary metabolites

Finally, 85 formulas could not be associated with known *Penicillium* secondary metabolites. This observation is in accord with the estimation based on genomic data claiming that more than 80% of fungal secondary metabolomes remain uncharacterized [11]. All detected secondary metabolites are listed in Table 1 with their structural hypotheses. Moreover, detailed lists of the major MS/MS fragment ions generated from each secondary metabolite by higher energy collision dissociation (HCD) (Table 2) as well as by collision-induced dissociation (CID) (Table S1, Supplementary data in the online version at DOI: [10.1016/j.jchromb.2017.03.011](https://doi.org/10.1016/j.jchromb.2017.03.011)) are provided together with their relative intensity. This complementary information is presented as an argument for further studies into these secondary metabolites. CID and HCD modes provide complementary information useful for structural analyses of fungal secondary metabolites. The resonant excitation mode used in CID within the linear ion trap rarely leads to consecutive fragmentations, whereas the non-resonant excitation of HCD could generate this kind of informative fragmentations. On the other hand, the CID mode is useful to study the filiation between the fragment ions formed in the consecutive MS<sup>n</sup> experiments. Unlike CID, the HCD mode is carried out without the low mass cut-off restriction, allowing the detection of low mass fragment ions.

In Table 1, several metabolites, including those identified, present the same chemical formula but are detected at different retention times. Since two isomers may display very different structures, MS/MS experiments were carried out. Two metabolites with the chemical formula  $C_{18}H_{20}N_4O_3$  were detected in the positive ionization mode (Table 1: PVer\_341.16.20.1 and PVer\_341.16.15.1). This formula was annotated as aurantiomide C (compound 10 in Fig. 3), and the two isomers displayed very similar MS/MS spectra (three same fragment ions in similar relative abundances, Table 2). Similarly, two isomers of  $C_{18}H_{22}N_4O_3$  with very similar MS/MS spectra and UV spectra (*data not shown*) (Table 2) were detected

in positive ionization mode and annotated as anacine (compound 8 in Fig. 3) (Table 1: PVer\_343.17.12.7 and PVer\_343.17.17.9).

Conversely, several metabolites with different molecular weights may share a common structure. For example,  $C_6H_{10}O_4$ , annotated by Antibase as verrucolone (PVer\_147.06.3.1), and  $C_6H_8O_3$  (PVer\_129.05.7.6) shared three common ions observed in their MS/MS spectra (Table 2), including two specific fragmentations:  $m/z=97.0280$  and  $m/z=87.0437$ . The metabolite  $C_6H_8O_3$  (PVer\_129.05.7.6) was therefore suspected to be a dehydroxylated form of verrucolone.

To overcome the complex manual comparison of hundreds of fragmentation patterns, some bioinformatics tools are available, such as MS2LDA [51] or the GNPS Molecular Networking system [34] which was developed for the study of natural products. This latter system highlights ions according to the similarity of their MS/MS spectra. In this way, a molecular network was generated from the MS/MS spectra of all annotated secondary metabolites (Fig. 5). This network displays already detected similar compounds such as ochratoxins A and B, or isomers, and ions linked through non-specific neutral losses:  $H_2O$  between ochratoxin B and the node 299.09 or  $NH_3$  and  $CO$  between verrucines and the node 391.14. The other groups consist of unknown molecules with the same formula.

More interestingly, the two metabolites at  $m/z=511.29$  shared seven common ions, suggesting a high similarity. The corresponding fragmentation patterns were scrutinized to explain the structural differences between the compound at  $R_T=18.0$ , annotated as the hydrolyzed fungisporin A, and the metabolite at  $R_T=18.7$ . The MS/MS spectra of the two metabolites are shown in Fig. 6. MS<sup>3</sup> and MS<sup>4</sup> experiments were performed to validate that the common ions displayed the same structure, and to obtain more structural information (Fig. S2 in the online version at DOI: [10.1016/j.jchromb.2017.03.011](https://doi.org/10.1016/j.jchromb.2017.03.011)). Both the MS/MS fragmentation patterns of the two metabolites show a dipole ion at  $m/z$  247.1435 and  $m/z$  265.1539 and consecutive losses of 99 and 147. Thus, it was suggested that  $m/z=247.1435$  consisted of a valine and a phenylalanine residue at the N-terminal extremity (Fig. 6, ions  $b_2$ ) and that the  $m/z$  265.1539 consisted of a valine and a phenylalanine residue in the C-terminal extremity of the two peptides (Fig. 6, ions  $y_2$ ). Moreover, the MS<sup>3</sup> spectrum of the  $m/z$  265 fragment ion of hydrolyzed fungisporin A showed a  $m/z$  118.0860 ( $C_5H_{12}NO_2$ ) corresponding to a C-terminal valine (Fig. S2A in the online version at DOI: [10.1016/j.jchromb.2017.03.011](https://doi.org/10.1016/j.jchromb.2017.03.011)), whereas the same MS<sup>3</sup> of PVer\_511.29.18.7 displayed a  $m/z$  166.0860 suggesting a C-terminal phenylalanine (Fig. S2B in the online version at DOI: [10.1016/j.jchromb.2017.03.011](https://doi.org/10.1016/j.jchromb.2017.03.011)). This was confirmed by the presence of a fragment ion at  $m/z$  346.2113 on its MS<sup>2</sup> spectrum (Fig. 6, ion  $b_3$ ). Finally, the MS<sup>3</sup> spectra of the  $m/z$  247.1435 fragment ion of both hydrolyzed fungisporin A and PVer\_511.29.18.7 were perfectly similar, suggesting that the N-terminal extremities of the two metabolites are identical. In this way, the linear tetrapeptide Val-Phe-Phe-Val was confirmed at  $R_T=18.0$  min. According to Ali et al., this molecule is suspected to result from the hydrolysis of a cyclic tetrapeptide, namely fungisporin A. Furthermore, structural analyses allowed the identification of its isomer detected at  $R_T=18.7$  min as the tetrapeptide Val-Phe-Val-Phe. Based on the assumption that this linear compound also resulted from the hydrolysis of a cyclic peptide [50], a cyclic Val-Phe-Val-Phe structure was suspected.

### 3.5. Conclusion

The use of two isotope labelings of fungal substrate and HPLC-HRMS analyses enabled the detection of 98 secondary metabolites produced by *P. verrucosum* grown on wheat grains, and led to the unambiguous determination of all their chemical formulas. Among these metabolites, 18 structural hypotheses were proposed from databases, at least nine of which were observed

**Table 2**  
Secondary metabolites detected from *Penicillium verrucosum* after culture on wheat grains and their MS/MS fragmentation in higher energy collision dissociation (HCD) mode. The fragmentation spectra are presented as a list of fragment ions with their relative intensity.

Identifier <sup>b</sup>	molecular formula <sup>a</sup>	parent ion	<sup>12</sup> C m/z (Da)	R <sub>T</sub> (min)	MS/MS mode (Normalized Collision Energy)	m/z of majors fragment ions (% relative intensity)
PVer_129.05.7.6	C <sub>6</sub> H <sub>8</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	129.0544	7.6	HCD (20)	129.0544(100) 87.0437(12) 97.0281(8) 55.0176(6) 111.0439(3)
PVer_147.06.3.1	C <sub>6</sub> H <sub>10</sub> O <sub>4</sub>	[M+H] <sup>+</sup>	147.0655	3.1	HCD (20)	147.0655(5) 129.0543(100) 117.0543(32) 97.0280(17) 85.0280(17) 99.0437(7) 87.0437(7) 75.0437(6) 55.0176(5) 71.0488(5) 115.0387(3)
PVer_231.12.5.6	C <sub>11</sub> H <sub>18</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	231.1229	5.6	HCD (20)	231.1227(100) 181.0857(70) 127.0387(27) 213.1120(26) 195.1014(25) 81.0332(22) 139.0750(21) 155.1064(21) 141.0543(13) 151.1116(10) 235.0616(100) 217.0510(85) 191.0718(70) 163.0770(5) 189.0562(2)
PVer_235.06.24.2	C <sub>12</sub> H <sub>12</sub> O <sub>5</sub>	[M–H] <sup>–</sup>	235.0616	24.2	HCD (60)	235.0616(100) 217.0510(85) 191.0718(70) 163.0770(5) 189.0562(2)
PVer_248.09.11.4	C <sub>13</sub> H <sub>13</sub> NO <sub>4</sub>	[M+H] <sup>+</sup>	248.0916	11.4	HCD (20)	248.0914(100) 230.0810(87)
PVer_235.10.8.3	C <sub>13</sub> H <sub>14</sub> O <sub>4</sub>	[M+H] <sup>+</sup>	235.0965	8.3	HCD (20)	235.0964(100) 217.0687(2) 191.0699(2)
PVer_235.10.19.8	C <sub>13</sub> H <sub>14</sub> O <sub>4</sub>	[M+H] <sup>+</sup>	235.0965	19.8	HCD (20)	235.0963(47) 193.0855(100) 165.0907(7) 217.0854(6) 207.1014(4) 139.1114(4) 211.0958(2) 137.0955(3)
PVer_251.09.23.5	C <sub>13</sub> H <sub>14</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	251.0912	23.5	HCD (20)	251.0915(100) 233.0807(45)
PVer_295.09.4.6	C <sub>13</sub> H <sub>16</sub> N <sub>2</sub> O <sub>6</sub>	[M–H] <sup>–</sup>	295.0936	4.6	HCD (60)	295.0936(15) 221.0573(100) 251.1039(20) 236.0807(7) 193.0622(25) 249.0885(6) 122.0252(9) 137.0361(9) 208.0859(8)
PVer_251.09.40.0	C <sub>13</sub> H <sub>16</sub> O <sub>5</sub>	[M–H] <sup>–</sup>	251.0924	40.0	HCD (60)	251.0924(100) 233.0823(9) 163.0769(9) 189.0562(5) 207.1030(6)
PVer_267.09.9.2	C <sub>13</sub> H <sub>16</sub> O <sub>6</sub>	[M–H] <sup>–</sup>	267.0871	9.2	HCD (60)	267.0871(100) 249.0772(6) 207.0666(16) 163.0769(5) 193.0875(4) 237.0773(3) 205.0875(2)
PVer_267.09.6.5	C <sub>13</sub> H <sub>16</sub> O <sub>6</sub>	[M–H] <sup>–</sup>	267.0871	6.5	HCD (60)	267.0871(100) 239.0721(5) 249.0563(1)
PVer_267.09.21.9	C <sub>13</sub> H <sub>16</sub> O <sub>6</sub>	[M–H] <sup>–</sup>	267.0871	21.9	HCD (60)	267.0871(95) 221.0825(100) 249.0773(10) 177.0926(10) 203.0719(4)
PVer_257.14.13.6	C <sub>13</sub> H <sub>20</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	257.1381	13.6	HCD (35)	257.1382(65) 157.0493(100)
PVer_299.08.3.5	C <sub>14</sub> H <sub>12</sub> N <sub>4</sub> O <sub>4</sub>	[M–H] <sup>–</sup>	299.0783	3.5	HCD (60)	299.0783(35) 256.0731(100) 228.0421(15) 145.0413(30) 213.0673(5) 238.0627(3) 185.0361(3)
PVer_238.06.11.5	C <sub>15</sub> H <sub>11</sub> O <sub>3</sub>	[M–H] <sup>–</sup>	238.0640	11.5	HCD (60)	238.0637(100) 210.0694(2)
PVer_242.16.16.1	C <sub>15</sub> H <sub>19</sub> N <sub>3</sub>	[M+H] <sup>+</sup>	242.1647	16.1	HCD (20)	242.650(100) 111.0915(33)
PVer_251.07.7.9	C <sub>16</sub> H <sub>12</sub> O <sub>3</sub>	[M–H] <sup>–</sup>	251.0712	7.9	HCD (60)	251.0712(100) 250.0646(13) 223.0771(6) 211.0405(4)
PVer_251.07.11.6	C <sub>16</sub> H <sub>12</sub> O <sub>3</sub>	[M–H] <sup>–</sup>	251.0712	11.6	HCD (60)	251.0712(100) 233.0612(3) 223.0770(2)
PVer_284.14.22.8	C <sub>16</sub> H <sub>17</sub> N <sub>3</sub> O <sub>2</sub>	[M+H] <sup>+</sup>	284.1388	22.8	HCD (20)	284.1390(100) 130.0648(30) 170.0597(3) 132.0804(3)
PVer_299.09.11.5	C <sub>17</sub> H <sub>14</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	299.0907	11.5	HCD (20)	299.0912(60) 205.0493(100) 281.0807(6) 189.0543(6) 253.0857(3) 107.0490(3) 187.0388(3) 159.0438(1) 263.0702(1)
PVer_297.08.11.5		[M–H] <sup>–</sup>	297.0773		HCD (60)	297.0773(10) 253.08874(100) 238.0641(6) 159.0456(4) 109.0300(4) 93.0350(4) 226.0626(2)
PVer_317.23.41.0	C <sub>17</sub> H <sub>34</sub> O <sub>5</sub>	[M–H] <sup>–</sup>	317.2327	41.0	HCD (60)	317.2327(100) 257.2127(52) 299.2229(6) 115.0770(3) 59.0143(18) 183.1395(4) 195.1759(4) 255.2334(4) 85.0663(1)
PVer_289.10.24.4	C <sub>18</sub> H <sub>14</sub> N <sub>2</sub> O <sub>2</sub>	[M–H] <sup>–</sup>	289.0977	24.4	HCD (60)	289.0977(100) 116.0510(33) 261.1039(24)
PVer_304.11.36.5	C <sub>18</sub> H <sub>15</sub> N <sub>3</sub> O <sub>2</sub>	[M–H] <sup>–</sup>	304.1087	36.5	HCD (60)	304.1087(26) 213.0549(100) 261.1037(29) 183.0567(8) 286.0988(6) 159.0568(3) 145.0410(3) 212.0470(3) 262.0989(3)
PVer_304.11.38.1	C <sub>18</sub> H <sub>15</sub> N <sub>3</sub> O <sub>2</sub>	[M–H] <sup>–</sup>	304.1087	38.1	HCD (60)	304.1087(100) 213.0548(100) 261.1038(25) 286.0998(5) 183.0569(7)
PVer_304.11.36.4	C <sub>18</sub> H <sub>15</sub> N <sub>3</sub> O <sub>2</sub>	[M–H] <sup>–</sup>	304.1087	36.4	HCD (60)	304.1102(3) 213.0548(100) 261.1040(25) 286.0992(5) 183.0569(7) 159.0569(4)
PVer_387.06.3.8	C <sub>18</sub> H <sub>16</sub> N <sub>2</sub> O <sub>6</sub> S	[M–H] <sup>–</sup>	387.0655	3.8	HCD (60)	387.0662(100) 176.0356(75) 353.0784(55) 202.0149(50) 269.0570(52) 121.0296(20) 231.0414(16) 277.0297(15) 121.0296(21) 270.0646(16)
PVer_341.16.20.1	C <sub>18</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	341.1604	20.1	HCD (20)	341.1605(55) 324.1338(100) 296.1389(10)
PVer_341.16.15.1	C <sub>18</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	341.1604	15.1	HCD (20)	341.1601(60) 324.1337(100) 296.1388(8) 151.1114(2) 195.1013(1)
PVer_343.17.12.7	C <sub>18</sub> H <sub>22</sub> N <sub>4</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	343.1747	12.7	HCD (20)	343.1758(45) 326.1496(100) 298.1546(5)
PVer_343.17.17.9	C <sub>18</sub> H <sub>22</sub> N <sub>4</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	343.1747	17.9	HCD (20)	343.1758(45) 326.1495(100) 298.1545(5) 270.0869(2)
PVer_351.14.18.1	C <sub>18</sub> H <sub>24</sub> O <sub>7</sub>	[M–H] <sup>–</sup>	351.1448	18.1	HCD (60)	351.1448(100) 307.1190(60) 249.0773(60) 333.1351(9) 205.0874(11) 101.0248(4) 219.1028(3)
PVer_351.14.29.0	C <sub>18</sub> H <sub>24</sub> O <sub>7</sub>	[M–H] <sup>–</sup>	351.1448	29.0	HCD (60)	351.1450(65) 201.0924(100) 245.1188(80) 277.1449(42) 263.1293(21) 231.1033(21) 233.1188(19) 213.0923(6)
PVer_351.14.24.6	C <sub>18</sub> H <sub>24</sub> O <sub>7</sub>	[M–H] <sup>–</sup>	351.1448	24.6	HCD (60)	351.1448(55) 321.1347(100) 307.1190(55) 245.0825(40) 201.0924(35) 233.1186(5) 259.0980(5)
PVer_351.14.29.1	C <sub>18</sub> H <sub>24</sub> O <sub>7</sub>	[M–H] <sup>–</sup>	351.1448	29.1	HCD (60)	351.1448(60) 201.0925(100) 245.1189(80) 277.1448(45)
PVer_351.14.24.1	C <sub>18</sub> H <sub>24</sub> O <sub>7</sub>	[M–H] <sup>–</sup>	351.1448	24.1	HCD (60)	351.1450(100) 321.1349(45) 249.0773(30) 307.1191(25) 245.0826(18) 201.0924(17) 101.0614(19)
PVer_277.21.39.4	C <sub>18</sub> H <sub>28</sub> O <sub>2</sub>	[M+H] <sup>+</sup>	277.2158	39.4	HCD (20)	277.2164(100) 179.1431(12) 135.1168(5) 259.2060(3) 121.1011(3) 235.1696(1) 161.1325(1)

Table 2 (Continued)

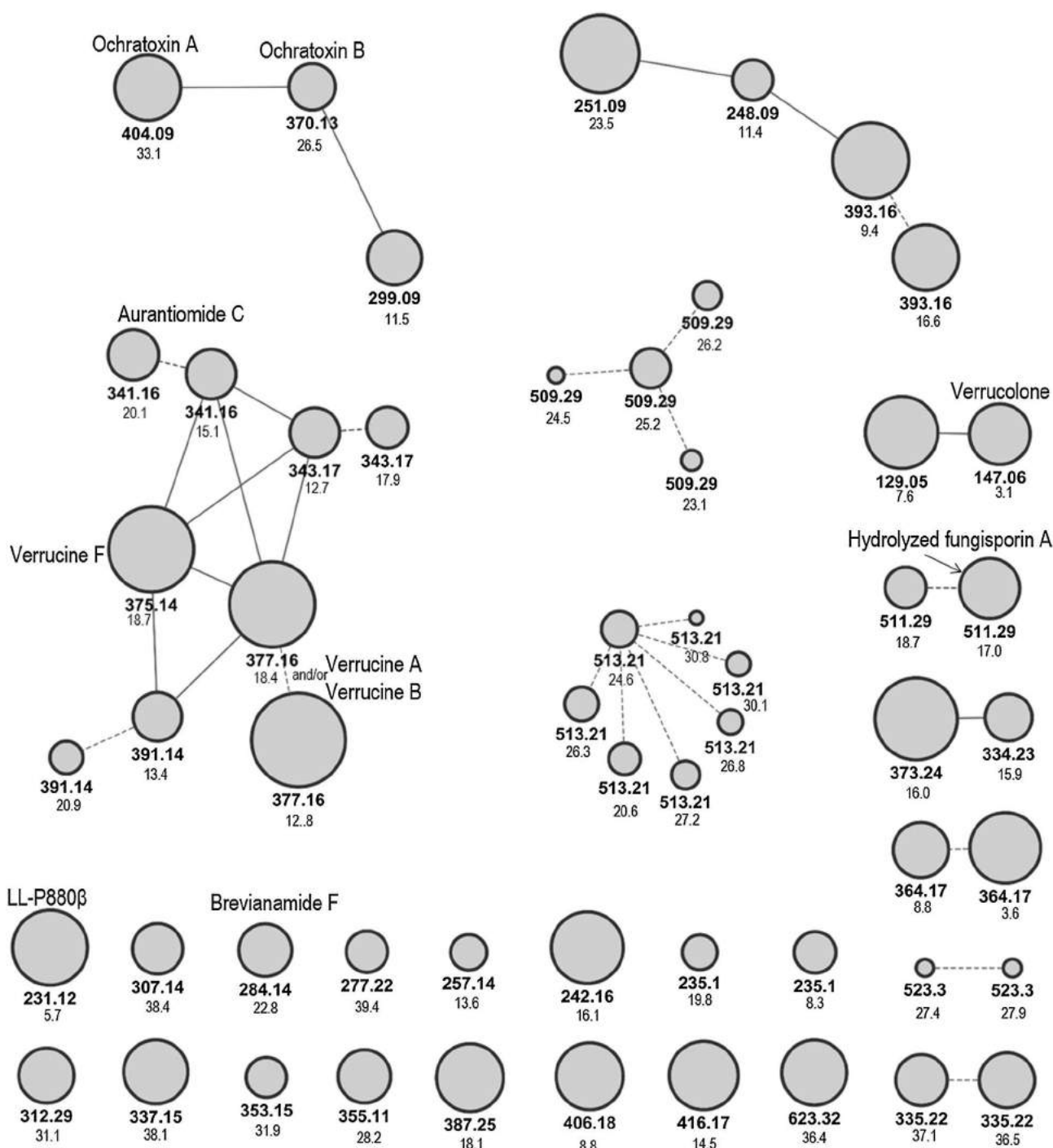
Identifier <sup>b</sup>	molecular formula <sup>a</sup>	parent ion	<sup>12</sup> C m/z (Da)	R <sub>T</sub> (min)	MS/MS mode (Normalized Collision Energy)	m/z of majors fragment ions (% relative intensity)
PVer_335.22.36.2	C <sub>18</sub> H <sub>32</sub> O <sub>4</sub>	[M+Na] <sup>+</sup>	335.2187	36.2	HCD (35)	335.2187(100) 317.2079(3) 190.1235(3)
PVer_335.22.36.8	C <sub>18</sub> H <sub>32</sub> O <sub>4</sub>	[M+Na] <sup>+</sup>	335.2187	36.8	HCD (35)	335.2187(100) 190.1232(20)
PVer_307.14.38.5	C <sub>19</sub> H <sub>18</sub> N <sub>2</sub> O <sub>2</sub>	[M+H] <sup>+</sup>	307.1427	38.5	HCD(35)	307.1439(75) 229.0971(100) 201.1022(12) 111.0553(10) 292.1207(2)
PVer_355.11.28.1	C <sub>19</sub> H <sub>18</sub> N <sub>2</sub> O <sub>3</sub> S	[M+H] <sup>+</sup>	355.1105	28.1	HCD (20)	355.1135(10) 289.0967(100) 337.1000(45) 307.1074(43) 308.1150(32) 279.1125(15) 290.1047(12) 236.1068(11) 261.1019(10) 327.1156(10) 371.1466(100) 353.1375(26)
PVer_371.15.22.5	C <sub>19</sub> H <sub>26</sub> O <sub>6</sub>	[M-2H+Na] <sup>-</sup>	371.1466	22.5	HCD (60)	371.1466(100) 353.1375(26)
PVer_312.29.31.9	C <sub>19</sub> H <sub>37</sub> NO <sub>2</sub>	[M+H] <sup>+</sup>	312.2895	31.9	HCD(20)	312.2895(5) 294.2792(100) 276.2684(5) 86.0598(5) 259.2417(5) 277.2524(1) 264.2683(1)
PVer_404.09.33.1	C <sub>20</sub> H <sub>18</sub> NO <sub>6</sub> Cl	[M+H] <sup>+</sup>	404.0897	33.1	HCD (20)	404.0892(30) 358.0836(100) 239.0103(66) 341.0572(13) 257.0209(11) 120.0804(10) 386.0786(5)
PVer_402.07.33.1		[M-H] <sup>-</sup>	402.0750		HCD (60)	402.0756(10) 358.0855(100) 211.0171(42) 254.0229(12) 239.0122(6) 322.1090(6) 314.0587(6) 270.0696(4)
PVer_370.13.26.5	C <sub>20</sub> H <sub>19</sub> NO <sub>6</sub>	[M+H] <sup>+</sup>	370.1285	26.5	HCD (20)	370.1286(30) 205.0492(100) 324.1226(37) 352.1172(7) 223.0598(7) 307.0959(6) 120.0805(6)
PVer_368.11.26.5		[M-H] <sup>-</sup>	368.1130		HCD (60)	368.1146(100) 324.1245(100) 220.0622(15) 177.0561(7) 280.0981(6) 164.0720(4)
PVer_337.15.38.1	C <sub>20</sub> H <sub>20</sub> N <sub>2</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	337.1552	38.1	HCD (20)	337.1552(100) 306.1360(54) 336.1952(10) 305.1281(7) 277.1331(6) 167.0851(3) 91.0539(4)
PVer_353.15.32.0	C <sub>20</sub> H <sub>20</sub> N <sub>2</sub> O <sub>4</sub>	[M+H] <sup>+</sup>	353.1480	32.0	HCD (20)	353.1489(100) 305.1278(56) 336.1460(7) 118.0409(4) 275.1020(2) 91.0538(1) 322.1309(1)
PVer_395.17.15.0	C <sub>20</sub> H <sub>28</sub> O <sub>8</sub>	[M-H] <sup>-</sup>	395.1708	15.0	HCD (60)	395.1708(46) 261.1866(100) 349.1661(62) 305.1762(60) 243.1760(36) 287.1656(29) 307.1919(25) 245.1553(17) 289.1813(9) 137.0612(7) 349.2021(9) 305.1761(100) 261.1865(15) 137.0612(7) 287.1656(5)
PVer_349.20.22.5	C <sub>20</sub> H <sub>30</sub> O <sub>5</sub>	[M-H] <sup>-</sup>	349.2012	22.5	HCD (60)	349.2021(9) 305.1761(100) 261.1865(15) 137.0612(7) 287.1656(5)
PVer_375.14.18.7	C <sub>21</sub> H <sub>18</sub> N <sub>4</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	375.1450	18.7	HCD (20)	375.1449(42) 358.1183(100) 330.1231(11)
PVer_373.13.17.0	C <sub>21</sub> H <sub>18</sub> N <sub>4</sub> O <sub>3</sub>	[M-H] <sup>-</sup>	373.1295	17.0	HCD (60)	373.1304(62) 302.0940(100) 262.0991(8) 145.0413(6) 355.1206(3)
PVer_391.14.13.5	C <sub>21</sub> H <sub>18</sub> N <sub>4</sub> O <sub>4</sub>	[M+H] <sup>+</sup>	391.1392	13.5	HCD (20)	391.1398(55) 374.1130(100) 346.1182(15)
PVer_389.12.13.5		[M-H] <sup>-</sup>	389.1258		HCD (60)	389.1264(100) 318.0897(99) 83.0145(15) 278.0948(7) 145.0417(5)
PVer_391.14.21.0	C <sub>21</sub> H <sub>18</sub> N <sub>4</sub> O <sub>4</sub>	[M+H] <sup>+</sup>	391.1392	21.0	HCD (20)	391.1396(60) 374.1129(100) 346.1182(10) 149.0230(9) 328.1078(1) 263.0810(1)
PVer_377.16.18.4	C <sub>21</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	377.1604	18.4	HCD (20)	377.1603(45) 360.1336(100) 332.1392(5) 91.0539(1)
PVer_377.16.12.8	C <sub>21</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	377.1604	12.8	HCD (20)	377.1603(43) 360.1339(100) 332.1389(5)
PVer_393.15.9.4	C <sub>21</sub> H <sub>20</sub> N <sub>4</sub> O <sub>4</sub>	[M+H] <sup>+</sup>	393.1554	9.4	HCD (20)	393.1551(18) 358.1181(100) 375.1448(75) 376.1284(55) 330.1227(7) 120.0441(6) 274.1180(3)
PVer_391.14.9.4		[M-H] <sup>-</sup>	391.1400		HCD(60)	391.1154(100) 250.1088(47) 294.0985(37) 373.1052(22) 300.0863(17) 276.0863(12) 302.0943(12)
PVer_393.15.15.5	C <sub>21</sub> H <sub>20</sub> N <sub>4</sub> O <sub>4</sub>	[M+H] <sup>+</sup>	393.1554	15.5	HCD (20)	393.1541(3) 358.1181(100) 375.1447(65) 330.1230(13) 376.1286(7) 348.1341(1)
PVer_364.16.3.6	C <sub>21</sub> H <sub>21</sub> N <sub>3</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	364.1656	3.6	HCD (20)	364.1649(100) 308.1026(10) 200.1066(6) 346.1544(3) 196.1115(2)
PVer_364.16.8.8	C <sub>21</sub> H <sub>21</sub> N <sub>3</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	364.1656	8.8	HCD (20)	364.1650(100) 308.1027(20) 177.0654(3) 200.1067(2) 294.1232(1) 346.1543(1) 145.1280(1)
PVer_401.22.25.8	C <sub>21</sub> H <sub>30</sub> N <sub>4</sub> O <sub>4</sub>	[M-H] <sup>-</sup>	401.2185	25.8	HCD (60)	401.2200(100) 262.1562(35) 172.1094(62) 214.1565(15) 245.1300(13) 228.1036(10) 198.0888(10) 203.0825(4) 137.0724(9) 290.1066(7)
PVer_401.22.21.9	C <sub>21</sub> H <sub>30</sub> N <sub>4</sub> O <sub>4</sub>	[M-H] <sup>-</sup>	401.2185	21.9	HCD (60)	401.2200(100) 262.1562(35) 172.1094(60) 214.1565(25) 245.1301(20) 203.0833(20) 155.0830(20) 137.0724(18) 228.1033(15) 198.0890(16) 146.0617(16)
PVer_334.23.15.8	C <sub>22</sub> H <sub>27</sub> N <sub>3</sub>	[M+H] <sup>+</sup>	334.2271	15.8	HCD (20)	334.2274(100) 203.1540(55) 217.1696(1)
PVer_332.21.35.6	C <sub>22</sub> H <sub>27</sub> N <sub>3</sub>	[M-H] <sup>-</sup>	332.2126	35.6	HCD (60)	332.2126(100) 264.1148(10) 240.1504(3)
PVer_416.17.14.6	C <sub>23</sub> H <sub>21</sub> N <sub>5</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	416.1712	14.6	HCD (20)	416.1709(40) 399.1445(100) 130.0648(56) 270.0870(45) 287.1136(40) 242.0919(1)
PVer_406.17.8.8	C <sub>23</sub> H <sub>23</sub> N <sub>3</sub> O <sub>4</sub>	[M+H] <sup>+</sup>	406.1758	8.8	HCD (20)	406.1753(21) 346.1544(100) 364.1649(62) 306.1232(19) 308.1027(10) 290.0921(4)
PVer_373.24.16.0	C <sub>24</sub> H <sub>28</sub> N <sub>4</sub>	[M+H] <sup>+</sup>	373.2385	16.0	HCD (20)	373.2381(100) 242.1649(55) 111.0914(7) 256.1802(1)
PVer_387.25.18.0	C <sub>25</sub> H <sub>30</sub> N <sub>4</sub>	[M+H] <sup>+</sup>	387.2536	18.0	HCD (50)	387.2542(3) 130.0652(100) 187.1232(10) 156.0810(10) 72.0807(7) 144.0809(6) 113.1074(4) 257.1889(2) 58.0650(2)

Table 2 (Continued)

Identifier <sup>b</sup>	molecular formula <sup>a</sup>	parent ion	<sup>12</sup> C <i>m/z</i> (Da)	R <sub>T</sub> (min)	MS/MS mode (Normalized Collision Energy)	<i>m/z</i> of majors fragment ions (% relative intensity)
PVer_385.24.35.7	C <sub>25</sub> H <sub>30</sub> N <sub>4</sub>	[M–H] <sup>–</sup>	385.2391	35.7	HCD (60)	385.2404(99) 230.1668(100) 256.2821(65) 116.0508(24) 156.0823(16)
PVer_385.24.37.2	C <sub>25</sub> H <sub>30</sub> N <sub>4</sub>	[M–H] <sup>–</sup>	385.2391	37.2	HCD (60)	385.2391(85) 230.1667(100) 256.1822(72) 116.0511(27) 156.0824(25) 142.0667(10)
PVer_513.21.20.6	C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	[M+Na] <sup>+</sup>	513.2119	20.6	HCD (20)	513.2129(60) 496.1858(100) 227.1063(25) 351.1674(4)
PVer_513.21.24.6	C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	[M+Na] <sup>+</sup>	513.2119	24.6	HCD (20)	513.2129(45) 496.1856(100) 468.1894(6)
PVer_513.21.26.3	C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	[M+Na] <sup>+</sup>	513.2119	26.3	HCD (20)	513.2129(100) 496.1858(65) 468.1930(4)
PVer_513.21.26.7	C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	[M+Na] <sup>+</sup>	513.2119	26.7	HCD (20)	513.2134(92) 496.1855(100)
PVer_513.21.27.1	C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	[M+Na] <sup>+</sup>	513.2119	27.1	HCD (20)	513.2128(63) 496.1852(100)
PVer_513.21.30.1	C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	[M+Na] <sup>+</sup>	513.2119	30.1	HCD (20)	513.2129(40) 496.1857(100) 351.1769(10)
PVer_513.21.31.4	C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	[M+Na] <sup>+</sup>	513.2119	31.4	HCD (20)	513.2134(55) 496.1863(100)
PVer_523.25.25.8	C <sub>28</sub> H <sub>36</sub> N <sub>4</sub> O <sub>6</sub>	[M–H] <sup>–</sup>	523.2555	25.8	HCD (60)	523.2562(5) 401.2204(100) 262.1565(16) 135.0455(16) 172.1097(10) 214.1570(4)
PVer_511.29.18.0	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	511.2909	18.0	HCD (20)	511.2906(10) 265.1542(100) 219.1488(30) 247.1438(21) 120.0804(16) 394.2120(10) 412.2228(9) 349.1902(3) 295.1438(3) 267.1484(2)
PVer_509.28.18.0		[M–H] <sup>–</sup>	509.2769		HCD (60)	509.2769(17) 116.0722(100) 276.1033(66) 115.0882(55) 202.1442(32) 146.0616(22) 263.1405(16) 164.0722(12) 262.1566(11) 246.1141(8) 174.0565(4)
PVer_511.29.18.7	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	511.2909	18.7	HCD (20)	511.2904(15) 265.1554(100) 219.1487(65) 247.1437(52) 346.2117(35) 120.0804(10) 72.0804(7) 166.0858(6) 301.1904(6)
PVer_509.28.18.7		[M–H] <sup>–</sup>	509.2769		HCD (60)	509.2769(16) 164.0722(100) 219.1506(20) 147.0457(19) 318.2191(10) 465.2877(15) 263.1404(6) 214.1564(7) 343.2139(6) 361.2249(5)
PVer_525.27.13.0	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>6</sub>	[M–H] <sup>–</sup>	525.2720	13.0	HCD (60)	525.2730(18) 116.0722(100) 292.0982(87) 172.1096(59) 278.1515(22) 162.0566(24) 263.1405(20) 115.0882(18) 248.1087(13) 142.0667(13) 202.1242(12)
PVer_509.29.26.2	C <sub>30</sub> H <sub>38</sub> N <sub>4</sub> O <sub>2</sub>	[M+Na] <sup>+</sup>	509.2894	26.2	HCD (20)	509.2899(40) 242.1651(100) 111.0915(20) 378.2171(12) 351.1741(6)
PVer_509.29.25.2	C <sub>30</sub> H <sub>38</sub> N <sub>4</sub> O <sub>2</sub>	[M+Na] <sup>+</sup>	509.2894	25.2	HCD (20)	509.2895(100) 242.1651(83) 111.0915(16) 378.2171(40)
PVer_509.29.23.0	C <sub>30</sub> H <sub>38</sub> N <sub>4</sub> O <sub>2</sub>	[M+Na] <sup>+</sup>	509.2894	23.0	HCD (20)	509.2895(73) 242.1347(100) 378.2168(13) 111.0913(13)
PVer_509.29.24.4	C <sub>30</sub> H <sub>38</sub> N <sub>4</sub> O <sub>2</sub>	[M+Na] <sup>+</sup>	509.2894	24.4	HCD (20)	509.2898(50) 242.1648(100) 378.2159(19) 111.0913(19) 478.0047(5)
PVer_548.29.17.4	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	[M–H] <sup>–</sup>	548.2892	17.4	HCD (60)	548.2892(10) 116.0722(100) 259.1457(65) 186.0565(45) 315.1144(30) 419.2307(20) 271.1252(15) 302.1515(15)
PVer_548.29.18.0	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	[M–H] <sup>–</sup>	548.2892	18.0	HCD (60)	548.2892(15) 116.0722(100) 172.1094(100) 315.1141(50) 258.1600(22) 271.1242(18) 186.0565(15) 419.2306(5)
PVer_548.29.19.1	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	[M–H] <sup>–</sup>	548.2892	19.1	HCD (60)	548.2892(12) 164.0722(100) 202.0990(40) 375.2409(32) 147.0457(18) 219.1505(10) 271.1780(8) 318.2193(8) 301.1675(7) 253.1677(5)
PVer_564.28.12.6	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	[M–H] <sup>–</sup>	564.2828	12.6	HCD (60)	553.2814(15) 116.0721(100) 202.0513(72) 285.1937(55) 119.0507(45) 331.1091(30) 435.2257(20) 302.1512(12) 278.1515(12) 155.0831(11) 241.1352(8)
PVer_567.09.26.3	C <sub>31</sub> H <sub>20</sub> O <sub>11</sub>	[M–H] <sup>–</sup>	567.0941	26.3	HCD (60)	567.0941(100) 523.1042(97) 257.0462(86) 265.0510(55) 309.0402(30) 256.0376(20)
PVer_567.09.25.1	C <sub>31</sub> H <sub>20</sub> O <sub>11</sub>	[M–H] <sup>–</sup>	567.0941	25.1	HCD (60)	567.0941(60) 523.1042(100) 439.0817(17)
PVer_523.30.27.7	C <sub>31</sub> H <sub>40</sub> N <sub>4</sub> O <sub>2</sub>	[M+Na] <sup>+</sup>	523.3047	27.7	HCD (20)	523.3038(12) 266.1168(100)
PVer_523.30.27.4	C <sub>31</sub> H <sub>40</sub> N <sub>4</sub> O <sub>2</sub>	[M+Na] <sup>+</sup>	523.3047	27.4	HCD (20)	523.3058(62) 266.1170(100)
PVer_507.28.38.5	C <sub>32</sub> H <sub>36</sub> N <sub>4</sub> O <sub>2</sub>	[M–H] <sup>–</sup>	507.2760	38.5	HCD (60)	507.2762(100) 236.0728(70)
PVer_537.29.39.5	C <sub>32</sub> H <sub>42</sub> O <sub>7</sub>	[M–H] <sup>–</sup>	537.2867	39.5	HCD (60)	537.2867(3) 71.0507(100) 95.0507(80) 147.0821(55) 159.1185(30) 417.2810(25) 321.2230(15) 399.2701(12)
PVer_553.28.38.7	C <sub>32</sub> H <sub>42</sub> O <sub>8</sub>	[M–H] <sup>–</sup>	553.2814	38.7	HCD (60)	553.2814(25) 175.1133(100) 121.0663(35) 97.0300(23) 253.1602(6)
PVer_553.28.38.3	C <sub>32</sub> H <sub>42</sub> O <sub>8</sub>	[M–H] <sup>–</sup>	553.2814	38.3	HCD (60)	553.2814(5) 175.1133(100) 121.0663(72) 97.0300(66) 201.0926(15) 393.1443(10) 135.0820(19) 119.0507(19)
PVer_623.32.36.3	C <sub>37</sub> H <sub>42</sub> N <sub>4</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	623.3215	36.3	HCD (20)	623.3226(2) 373.2386(100) 242.1653(25) 251.0916(13) 111.0916(3)
PVer_621.30.36.3		[M–H] <sup>–</sup>	621.3073		HCD (40)	621.3091(5) 371.2249(100) 249.0773(62)
PVer_713.28.33.4	C <sub>38</sub> H <sub>42</sub> N <sub>4</sub> O <sub>10</sub>	[M–H] <sup>–</sup>	713.2847	33.4	HCD (60)	713.2847(2) 190.0276(100) 191.0355(65) 309.0411(22) 161.0249(17) 311.0567(12) 325.0723(9) 400.2126(2)

<sup>a</sup> The chemical formulas were calculated after comparison of the *m/z* ratios obtained from the culture of *P. verrucosum* in <sup>12</sup>C wheat, 97% <sup>13</sup>C wheat and 53% <sup>13</sup>C/96% <sup>15</sup>N wheat.

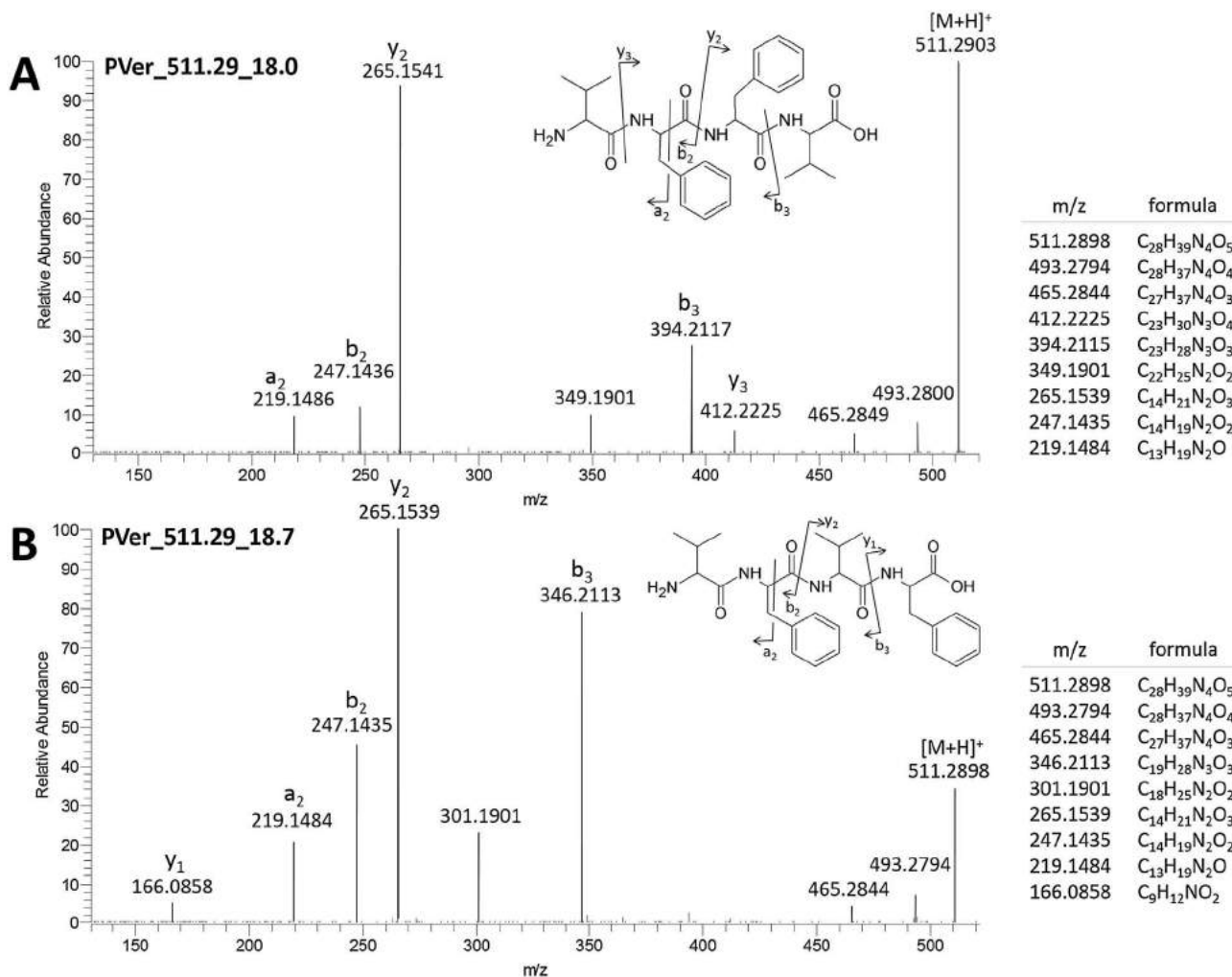
<sup>b</sup> Secondary metabolites identifier (PVer.*m/z*.R<sub>T</sub>).



**Fig. 5.** Molecular network of the secondary metabolome of *Penicillium verrucosum*, created with the GNPS software from the targeted MS/MS spectra (ESI<sup>+</sup> high-resolution CID): nodes name: *m/z* ratio and retention time of the precursor ion, node size: absolute intensity in a logarithmic scale of the precursor ion in the HRMS spectrum, black edges: connected MS/MS spectra according to GNPS software, red dotted edges: manually connected isomers which were not detected by GNPS.

for the first time in *P. verrucosum*: anacine, aurantioamide C, brevianamide F, LL-P880β and at least six fungisporins. The lack of available reference compounds remains problematic for natural product studies. This can be partially balanced by using informatics tools for the simulation of fragmentation spectra. In this study, CFM-ID was successfully used for several *P. verrucosum* metabolites, providing relevant supporting information about the metabolite annotations. However, 79 secondary metabolites remained unidentified. A molecular network was generated with the GNPS system, enabling the detection of a metabolite of interest. This new metabo-

lite was characterized by MS<sup>n</sup> experiments as a linear tetrapeptide, most likely resulting from the hydrolysis of the cyclic Val-Phe-Val-Phe. In any case, the 98 metabolites detected in this study constitute a valuable source of information for further works focused on identification of new metabolites. These data will be useful for purification and structural analyses with the help of complementary analyses by NMR for example, for application to the particular *P. verrucosum* metabolome, or more generally to other *Penicillium* metabolomes.



**Fig. 6.** Comparison of the MS/MS spectra (ESI<sup>+</sup> high-resolution CID) of two secondary metabolites produced by *Penicillium verrucosum* grown on wheat grains. (A) MS/MS spectra of C<sub>28</sub>H<sub>38</sub>N<sub>4</sub>O<sub>5</sub> detected at R<sub>T</sub> = 18.0 min (PVer\_511.29\_18.0). (B) MS/MS spectra of C<sub>28</sub>H<sub>38</sub>N<sub>4</sub>O<sub>5</sub> detected at R<sub>T</sub> = 18.7 min (PVer\_511.29\_18.7).

## Funding sources

This study was co-funded by INRA and French Minister of Higher Education and Research in the context of a project supported by French National Agency of Research (ANR-15-CE21-0010-21 Newmyco).

## Acknowledgements

The labeled substrates were produced by the Groupe de Recherches Appliquées en Phytotechnologie, CEA, IBEB, Cadarache, FR 13108 Saint-Paul-les-Durance, France. The authors are grateful to Andrew Morris who corrected the manuscript. The authors also are grateful to Pascal Bourdaudhui, who developed the software MassCompare, and to Jean-Claude Tabet for his kind help with the mass spectrometry analyses and data monitoring. The authors would also like to thank the French National Infrastructure of Metabolomics and Fluxomics (MetaboHUB-ANR-11-INBS-0010) for their support.

## References

- [1] A.A. Brakhage, Regulation of fungal secondary metabolism, *Nat. Rev. Microbiol.* 11 (2013) 21–32.

- [2] R. Hare, W. Florey, A.D. Gardner, N.G. Heatley, M.A. Jennings, J. Orr-Ewing, A.G. Sanders, E.P. Abraham, E. Chain, C.M. Fletcher, H.W. Florey, New light on the history of penicillin, *Med. Hist.* 26 (1982) 1–24.
- [3] A.A. Brakhage, V. Schroeckh, Fungal secondary metabolites – strategies to activate silent gene clusters, *Fungal Genet. Biol.* 48 (2011) 15–22.
- [4] C. Bömkje, B. Tudzynski, Diversity, regulation, and evolution of the gibberellin biosynthetic pathway in fungi compared to plants and bacteria, *Phytochemistry* 70 (2009) 1876–1893.
- [5] J. Avalos, A. Prado-Cabrero, A. Estrada, Neurosporaxanthin production by *Neurospora* and *Fusarium*, in: J.-L. Barredo (Ed.), *Microb. Carotenoids from Fungi Methods Protoc*, Springer S, New York, 2012, pp. 153–159, 2012.
- [6] T. Kuiper-Goodman, Toxicology letters risk assessment, *Toxicol. Lett.* 82/83 (1995) 853–859.
- [7] D.L. Eaton, E.P. Gallagher, Mechanisms of aflatoxin carcinogenesis, *Annu. Rev. Pharmacol. Toxicol.* 34 (1994) 135–172.
- [8] T.R. Bui-Klimke, F. Wu, Ochratoxin A and human health risk: a review of the evidence, *Crit. Rev. Food Sci. Nutr.* 55 (2015) 1860–1869.
- [9] A. Zinedine, J.M. Soriano, J.C. Moltó, J. Mañes, Review on the toxicity, occurrence, metabolism, detoxification, regulations and intake of zearalenone: an oestrogenic mycotoxin, *Food Chem. Toxicol.* 45 (2007) 1–18.
- [10] J.W. Bennett, M. Klich, Mycotoxins, *Clin. Microbiol. Rev.* 16 (2003) 497–516.
- [11] E.M. Fox, B.J. Howlett, Secondary metabolism: regulation and role in fungal biology, *Curr. Opin. Microbiol.* 11 (2008) 481–487.
- [12] S. Elmholt, P.H. Rasmussen, *Penicillium verrucosum* occurrence and ochratoxin A contents in organically cultivated grain with special reference to ancient wheat types and drying practice, *Mycopathologia* 159 (2005) 421–432.
- [13] N. Magan, R. Hope, V. Cairns, D. Aldred, Post-harvest fungal ecology: impact of fungal growth and mycotoxin accumulation in stored grain, *Eur. J. Plant Pathol.* 109 (2003) 723–730.
- [14] F. Wu, T. Bui-Klimke, K. Naumoff Shields, Potential economic and health impacts of ochratoxin A regulatory standards, *World Mycotoxin J.* 7 (2014) 387–398.

- [15] T. Kuiper-Goodman, P.M. Scott, Risk assessment of the mycotoxin ochratoxin A, *Biomed. Environ. Sci.* 2 (1989) 179–248.
- [16] International Agency for Research on Cancer (IARC), Ochratoxin A, IARC Monogr. Eval. Carcinog. Risk Chem. Hum. 56 (1993) 489–521.
- [17] P. Kroghl, B. Haidt, R. Pleftinas, S. Ceovic, Balkan (endemic) nephropathy and foodborn ochratoxin A: preliminary results of a survey of foodstuffs, *Acta Path. Microbiol. Scand. Sect. B* 85 (1977) 238–240.
- [18] F. Malir, V. Ostry, A. Pfohl-Leszakowicz, J. Malir, J. Toman, Ochratoxin A: 50 years of research, *Toxins (Basel)* 8 (2016) 12–15.
- [19] T. Bui-Klimke, F. Wu, Evaluating weight of evidence in the mystery of balkan endemic nephropathy, *Risk Anal.* 34 (2014) 1688–1705.
- [20] T.O. Larsen, A. Svendsen, J. Smedsgaard, Biochemical characterization of ochratoxin A-producing strains of the genus *Penicillium*, *Appl. Environ. Microbiol.* 67 (2001) 3630–3635.
- [21] M.T. Henke, N.L. Kelleher, Modern mass spectrometry for synthetic biology and structure-based discovery of natural products, *Nat. Prod. Rep.* 33 (2016) 942–950.
- [22] W.B. Dunn, A. Erban, R.J.M. Weber, D.J. Creek, M. Brown, R. Breitling, T. Hankemeier, R. Goodacre, S. Neumann, J. Kopka, M.R. Viant, Mass appeal: metabolite identification in mass spectrometry-focused untargeted metabolomics, *Metabolomics* 9 (2013) 44–66.
- [23] C. Bueschl, R. Krska, B. Kluger, R. Schuhmacher, Isotopic labeling-assisted metabolomics using LC-MS, *Anal. Bioanal. Chem.* 405 (2013) 27–33.
- [24] C. Bueschl, B. Kluger, M. Lemmens, G. Adam, G. Wiesenberger, V. Maschietto, A. Marocco, J. Strauss, S. Bödi, G.G. Thallinger, R. Krska, R. Schuhmacher, A novel stable isotope labelling assisted workflow for improved untargeted LC-HRMS based metabolomics research, *Metabolomics* 10 (2014) 754–769.
- [25] N.K.N. Neumann, S.M. Lehner, B. Kluger, C. Bueschl, K. Sedelmaier, M. Lemmens, R. Krska, R. Schuhmacher, Automated LC-HRMS(/MS) approach for the annotation of fragment ions derived from stable isotope labeling-assisted untargeted metabolomics, *Anal. Chem.* 86 (2014) 7320–7327.
- [26] H.B. Bode, B. Bethe, R. Hofs, A. Zeek, Big effects from small changes: possible ways to explore nature's chemical diversity, *ChemBioChem* 3 (2002) 619–627.
- [27] M. Hallas-Møller, K.F. Nielsen, J.C. Frisvad, Influence of the medium and preferred cereal substrate on secondary metabolite production by species from *Penicillium* series *Viridicata*, *Planta Med.* 82 (2016) S1–S381.
- [28] P.M. Cano, E.L. Jamin, S. Tadriss, P. Bourdaud'Hui, M. Péan, L. Debrauwer, I.P. Oswald, M. Delaforge, O. Puel, New untargeted metabolic profiling combining mass spectrometry and isotopic labeling: application on *Aspergillus fumigatus* grown on wheat, *Anal. Chem.* 85 (2013) 8412–8420.
- [29] J.C. Frisvad, R.A. Samson, Polyphasic taxonomy of *Penicillium* subgenus *Penicillium*: a guide to identification of food and air-borne terverticillate *Penicillia* and their mycotoxins, *Stud. Mycol.* 49 (2002) 1–173.
- [30] Y.C.S. Adjovi, S. Bailly, B.J.G. Gnonlonfin, S. Tadriss, A. Querin, A. Sanni, I.P. Oswald, O. Puel, J.D. Bailly, Analysis of the contrast between natural occurrence of toxigenic *Aspergilli* of the *Flavi* section and aflatoxin B1 in cassava, *Food Microbiol.* 38 (2014) 151–159.
- [31] M. Pean, S. Boiry, J.-C. Ferrandi, F. Gibiat, O. Puel, M. Delaforge, Production and use of mycotoxins uniformly enriched with stable isotopes for their dosage in biological samples: (1) production of uniformly enriched biomass, *J. Label. Compd. Radiopharm.* 50 (2007) 563–564.
- [32] L.W. Sumner, A. Amberg, D. Barrett, M.H. Beale, R. Beger, C.A. Daykin, T.W. Fan, O. Fiehn, R. Goodacre, J.L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, Proposed minimum reporting standards for chemical analysis chemical analysis working group (CAWG) metabolomics standards initiative (MSI), *NIH* 3 (2007) 211–221.
- [33] F. Allen, A. Pon, M. Wilson, R. Greiner, D. Wishart, CFM-ID: a web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra, *Nucleic Acids Res.* 42 (2014) 1–6.
- [34] M. Wang, J.J. Carver, V.V. Phelan, L.M. Sanchez, N. Garg, Y. Peng, D.D. Nguyen, J. Watrous, C.A. Kapon, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A.V. Melnik, M.J. Meehan, W.-T. Liu, M. Crüsemann, P.D. Boudreau, E. Esquenazi, M. Sandoval-Calderón, R.D. Kersten, L.A. Pace, R.A. Quinn, K.R. Duncan, C.-C. Hsu, D.J. Floros, R.G. Gavilan, K. Kleigrewe, T. Northen, R.J. Dutton, D. Parrot, E.E. Carlson, B. Aigle, C.F. Michelsen, L. Jelsbak, C. Sohlenkamp, P. Pevzner, A. Edlund, J. McLean, J. Piel, B.T. Murphy, L. Gerwick, C.-C. Liaw, Y.-L. Yang, H.-U. Humpf, M. Maansson, R.A. Keyzers, A.C. Sims, A.R. Johnson, A.M. Sidebottom, B.E. Sedio, A. Klitgaard, C.B. Larson, C.A. Boya, P. D. Torres-Mendoza, D.J. Gonzalez, D.B. Silva, L.M. Marques, D.P. Demarque, E. Pociute, E.C. O'Neill, E. Briand, E.J.N. Helfrich, E.A. Granatosky, E. Glukhov, F. Ryffel, H. Houson, H. Mohimani, J.J. Kharbush, Y. Zeng, J.A. Vorholt, K.L. Kurita, P. Charusanti, K.L. McPhail, K.F. Nielsen, L. Vuong, M. Elfeki, M.F. Traxler, N. Engene, N. Koyama, O.B. Vining, R. Baric, R.R. Silva, S.J. Mascuch, S. Tomasi, S. Jenkins, V. Macherla, T. Hoffman, V. Agarwal, P.G. Williams, J. Dai, R. Neupane, J. Gurr, A.M.C. Rodríguez, A. Lamsa, C. Zhang, K. Dorresteijn, B.M. Duggan, J. Almaliti, P.-M. Allard, P. Phapale, L.-F. Nothias, T. Alexandrov, M. Litaudon, J.-L. Wolfender, J.E. Kyle, T.O. Metz, T. Peryea, D.-T. Nguyen, D. VanLeer, P. Shinn, A. Jadhav, R. Müller, K.M. Waters, W. Shi, X. Liu, L. Zhang, R. Knight, P.R. Jensen, B.Ø. Palsson, K. Pogliano, R.G. Lington, M. Gutiérrez, N.P. Lopes, W.H. Gerwick, B.S. Moore, P.C. Dorresteijn, N. Bandeira, Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking, *Nat. Biotechnol.* 34 (2016) 828–837.
- [35] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Res.* 13 (2003) 2498–2504.
- [36] R.A. Samson, K.A. Seifert, A.F. Kuijpers, J.A.M. Houbraken, J.C. Frisvad, Phylogenetic analysis of *Penicillium* subgenus *Penicillium* using partial  $\beta$ -tubulin sequence, *Stud. Mycol.* 49 (2004) 175–200.
- [37] K.F. Nielsen, T.O. Larsen, The importance of mass spectrometric dereplication in fungal secondary metabolite analysis, *Front. Microbiol.* 6 (2015) 71.
- [38] Y.H. Chooi, Y. Tang, Navigating the fungal polyketide chemical space: from genes to molecules, *J. Org. Chem.* 77 (2012) 9933–9953.
- [39] H. Laatsch, *Antibase 2012: The Natural Compound Identifier*, 2012, <https://application.wiley-vch.de/stmdata/antibase.php> (Accessed 2 January 2017).
- [40] S.C. Duarte, A. Pena, C.M. Lino, A review on ochratoxin A occurrence and effects of processing of cereal and cereal derived food products, *Food Microbiol.* 27 (2010) 187–198.
- [41] J.C. Frisvad, O. Filtenborg, Classification of terverticillate *Penicillia* based on profiles of mycotoxins and other secondary metabolites, *Appl. Environ. Microbiol.* 46 (1983) 1301–1310.
- [42] M. Kokkonen, M. Jestoi, A. Rizzo, The effect of substrate on mycotoxin production of selected *Penicillium* strains, *Int. J. Food Microbiol.* 99 (2005) 207–214.
- [43] T.O. Larsen, H. Franzky, S.R. Jensen, UV-guided isolation of verrucines A and B, novel quinazolines from *Penicillium verrucosum* structurally related to anacine from *Penicillium aurantiogriseum*, *J. Nat. Prod.* 62 (1999) 1578–1580.
- [44] S.L. Leong, J. Schnürer, A. Broberg, F. Verrucine, a quinazoline from *Penicillium verrucosum*, *J. Nat. Prod.* 71 (2008) 1455–1457.
- [45] L. Rahbaek, S. Sperry, J.C. Frisvad, T.O. Larsen, PC-2 LL-P888gamma and some novel analogue alpha-pyrones from *Penicillium nordicum*, *P. verrucosum* and *P. olsonii*, *Biochem. Syst. Ecol.* 31 (2003) 313–317.
- [46] J.M. Boyes-Korkis, K.A. Gurney, J. Penn, P.G. Mantle, J.N. Bilton, R.N. Sheppard, Anacine, a new benzodiazepine metabolite of *Penicillium aurantiogriseum* produced with other alkaloids in submerged fermentation, *J. Nat. Prod.* 56 (1993) 1707–1717.
- [47] Z.H. Xin, Y. Fang, L. Du, T. Zhu, L. Duan, J. Chen, Q.Q. Gu, W.M. Zhu, Aurantiomides A–C, quinazoline alkaloids from the sponge-derived fungus *Penicillium aurantiogriseum* SP 0–19, *J. Nat. Prod.* 70 (2007) 853–855.
- [48] S. Maiya, A. Grundmann, S.M. Li, G. Turner, The fumitremorgin gene cluster of *Aspergillus fumigatus*: identification of a gene encoding brevianamide F synthetase, *ChemBioChem* 7 (2006) 1062–1069.
- [49] A. Klitgaard, J.B. Nielsen, R.J.N. Frandsen, M.R. Andersen, K.F. Nielsen, Combining stable isotope labeling and molecular networking for biosynthetic pathway characterization, *Anal. Chem.* 87 (2015) 6520–6526.
- [50] H. Ali, M.I. Ries, P.P. Lankhorst, R.A.M. Van Der Hoeven, O.L. Schouten, M. Noga, T. Hankemeier, N.N.M.E. Van Peij, R.A.L. Bovenberg, R.J. Vreeken, A.J.M. Driessen, A non-canonical NRPS is involved in the synthesis of fungisporin and related hydrophobic cyclic tetrapeptides in *Penicillium chrysogenum*, *PLoS One* 9 (2014) e98212.
- [51] J.J.J. van der Hooft, J. Wandy, M.P. Barrett, K.E.V. Burgess, S. Rogers, Topic modeling for untargeted substrate exploration in metabolomics, *Proc. Natl. Acad. Sci. U. S. A.* 113 (2016) 13738–13743.



SUPPLEMENTARY MATERIAL

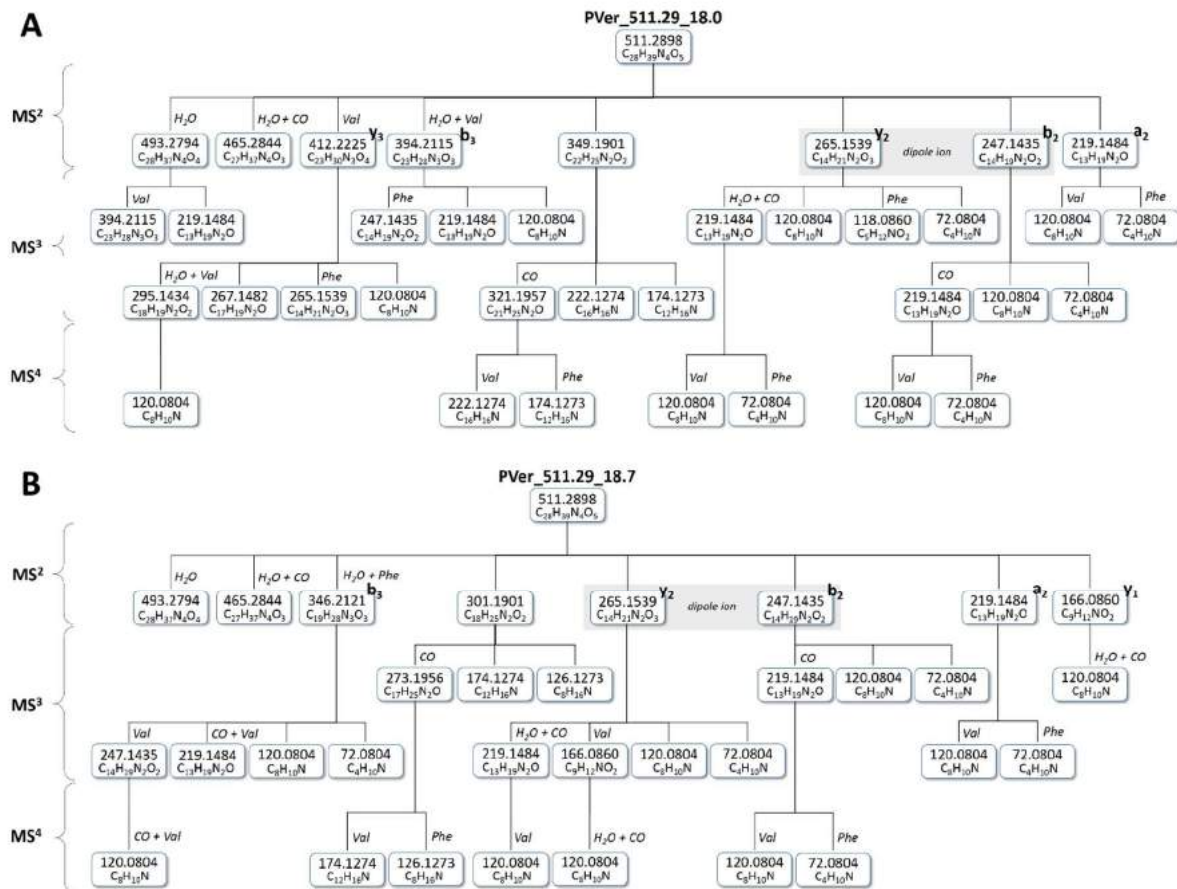


Figure S2. Fragmentation trees of two secondary metabolites produced by *Penicillium verrucosum*. (A) Fragmentation tree of C<sub>28</sub>H<sub>38</sub>N<sub>4</sub>O<sub>5</sub> detected at R<sub>T</sub>= 18.0 min (PVer\_511.29\_18.0). (B) Fragmentation tree of C<sub>28</sub>H<sub>38</sub>N<sub>4</sub>O<sub>5</sub> detected at R<sub>T</sub>= 18.7 min (PVer\_511.29\_18.7). MS<sup>2</sup> spectra were acquired in ESI+ high-resolution CID mode. MS<sup>3</sup> spectra were generated by CID fragmentation of the parent ion (MS<sup>2</sup>) and subsequent HCD fragmentation of the daughter ion. MS<sup>4</sup> experiments were performed after CID fragmentation of parent and daughter ions and HCD fragmentation of the third generation product ion.

**Travail expérimental** – Caractérisation du métabolome secondaire de *P. verrucosum*

molecular formula <sup>a</sup>	Identifier <sup>f</sup>	parent ion	<sup>12</sup> C m/z (Da)	R <sub>T</sub> (min)	MS/MS mode (Normalized Collision Energy)	m/z of majors fragment ions (% relative intensity)
C <sub>6</sub> H <sub>8</sub> O <sub>3</sub>	PVer_129.05_7.6	[M+H] <sup>+</sup>	129.0544	7.6	CID (25)	129.0544(100) 87.0437(7) 97.0281(18) 55.0176(7) 111.0439(7)
C <sub>6</sub> H <sub>10</sub> O <sub>4</sub>	PVer_147.06_3.1	[M+H] <sup>+</sup>	147.0655	3.1	CID (25)	147.0655(4) 129.0543(100) 117.0543(76) 97.0280(5) 85.0280(6) 87.0437(1) 75.0437(3) 115.0387(1) 119.0854(5)
C <sub>11</sub> H <sub>18</sub> O <sub>5</sub>	PVer_231.12_5.6	[M+H] <sup>+</sup>	231.1229	5.6	CID (25)	231.1227(45) 181.0857(100) 127.0387(72) 213.1120(42) 195.1014(23) 81.0332(23) 139.0750(20) 155.1064(42) 141.0543(37) 151.1116(20)
C <sub>12</sub> H <sub>12</sub> O <sub>5</sub>	PVer_235.06_24.2	[M-H] <sup>-</sup>	235.0616	24.2	CID (15)	235.0616(100) 217.0510(27) 191.0718(22)
C <sub>13</sub> H <sub>13</sub> NO <sub>4</sub>	PVer_248.09_11.4	[M+H] <sup>+</sup>	248.0916	11.4	CID (15)	248.0914(7) 230.0810(100)
C <sub>13</sub> H <sub>14</sub> O <sub>4</sub>	PVer_235.10_8.3	[M+H] <sup>+</sup>	235.0965	8.3	CID (25)	235.0964(2) 217.0687(10) 191.0699(100)
C <sub>13</sub> H <sub>14</sub> O <sub>4</sub>	PVer_235.10_19.8	[M+H] <sup>+</sup>	235.0965	19.8	CID(25)	235.0963(3) 193.0855(100) 165.0907(1) 217.0854(10) 207.1014(9) 139.1114(3) 204.0889(5) 191.0698(5) 179.1063(7)
C <sub>13</sub> H <sub>14</sub> O <sub>5</sub>	PVer_251.09_23.5	[M+H] <sup>+</sup>	251.0912	23.5	CID (25)	251.0915(12) 233.0807(100)
C <sub>13</sub> H <sub>16</sub> N <sub>2</sub> O <sub>6</sub>	PVer_295.09_4.6	[M-H] <sup>-</sup>	295.0936	4.6	CID (15)	295.0936(40) 221.0573(4) 251.1039(100) 236.0807(85) 249.0885(52) 251.0718(20)
C <sub>13</sub> H <sub>16</sub> O <sub>5</sub>	PVer_251.09_40.0	[M-H] <sup>-</sup>	251.0924	40.0	CID (15)	251.0924(100) 233.0823(17) 163.0769(6) 189.0562(2) 207.1030(4)
C <sub>13</sub> H <sub>16</sub> O <sub>6</sub>	PVer_267.09_9.2	[M-H] <sup>-</sup>	267.0871	9.2	CID (15)	267.0871(100) 249.0772(19) 207.0666(15) 237.0773(5) 205.0875(2) 193.0875(2) 163.0769(1) 239.0720(1) 179.0718(2)
C <sub>13</sub> H <sub>16</sub> O <sub>6</sub>	PVer_267.09_6.5	[M-H] <sup>-</sup>	267.0871	6.5	CID (15)	267.0871(100) 239.0721(4) 237.0774(5) 249.0772(2) 207.0666(2) 193.0875(2)
C <sub>13</sub> H <sub>16</sub> O <sub>6</sub>	PVer_267.09_21.9	[M-H] <sup>-</sup>	267.0871	21.9	CID (15)	267.0871(55) 221.0825(100) 249.0773(30) 249.0773(30) 239.0719(4) 179.0716(2) 207.0666(2)
C <sub>13</sub> H <sub>20</sub> O <sub>5</sub>	PVer_257.14_13.6	[M+H] <sup>+</sup>	257.1381	13.6	CID (15)	257.1382(90) 157.0493(100)
C <sub>14</sub> H <sub>12</sub> N <sub>4</sub> O <sub>4</sub>	PVer_299.08_3.5	[M-H] <sup>-</sup>	299.0783	3.5	CID (15)	299.0783(100) 256.0731(55) 228.0421(7)
C <sub>15</sub> H <sub>11</sub> O <sub>3</sub>	PVer_238.06_11.5	[M-H] <sup>-</sup>	238.0640	11.5	CID (15)	238.0637(100) 210.0694(2)
C <sub>15</sub> H <sub>19</sub> N <sub>3</sub>	PVer_242.16_16.1	[M+H] <sup>+</sup>	242.1647	16.1	CID (25)	242.650(27) 111.0915(100) 211.1229(7) 96.0679(4) 183.0914(2)
C <sub>16</sub> H <sub>12</sub> O <sub>3</sub>	PVer_251.07_7.9	[M-H] <sup>-</sup>	251.0712	7.9	CID (15)	251.0712(100) 250.0646(7) 223.0771(5) 211.0405(2) 233.0613(2)
C <sub>16</sub> H <sub>12</sub> O <sub>3</sub>	PVer_251.07_11.6	[M-H] <sup>-</sup>	251.0712	11.6	CID (15)	251.0712(100) 233.0612(4) 223.0770(2)
C <sub>16</sub> H <sub>17</sub> N <sub>3</sub> O <sub>2</sub>	PVer_284.14_22.8	[M+H] <sup>+</sup>	284.1388	22.8	CID (15)	284.1390(33) 130.0648(100) 170.0597(24) 132.0804(33) 153.0655(7) 239.1177(7) 267.1125(9) 115.0864(4) 256.1445(1)
C <sub>17</sub> H <sub>14</sub> O <sub>5</sub>	PVer_299.09_11.5	[M+H] <sup>+</sup>	299.0907	11.5	CID (25)	299.0912(4) 205.0493(100) 281.0807(7) 189.0543(6) 253.0857(1) 187.0388(4) 263.0702(1)
	PVer_297.08_11.5	[M-H] <sup>-</sup>	297.0773		CID (15)	297.0773(40) 253.08874(100) 159.0456(2) 226.0626(2)
C <sub>17</sub> H <sub>34</sub> O <sub>5</sub>	PVer_317.23_41.0	[M-H] <sup>-</sup>	317.2327	41.0	CID (15)	317.2327(100) 257.2127(15) 299.2229(6) 255.2331(2)
C <sub>18</sub> H <sub>14</sub> N <sub>2</sub> O <sub>2</sub>	PVer_289.10_24.4	[M-H] <sup>-</sup>	289.0977	24.4	CID (15)	289.0977(100) 116.0510(4) 261.1039(12) 271.0978(3) 245.0827(3)
C <sub>18</sub> H <sub>15</sub> N <sub>3</sub> O <sub>2</sub>	PVer_304.11_36.5	[M-H] <sup>-</sup>	304.1087	36.5	CID (15)	304.1087(100) 213.0549(37) 261.1037(37) 303.0940(5)
C <sub>18</sub> H <sub>15</sub> N <sub>3</sub> O <sub>2</sub>	PVer_304.11_38.1	[M-H] <sup>-</sup>	304.1087	38.1	CID (15)	304.1087(45) 213.0548(100) 261.1038(99) 286.0998(4) 276.0935(8)
C <sub>18</sub> H <sub>15</sub> N <sub>3</sub> O <sub>2</sub>	PVer_304.11_36.4	[M-H] <sup>-</sup>	304.1087	36.4	CID (15)	304.1102(100) 213.0548(77) 261.1040(75) 286.0992(5)
C <sub>18</sub> H <sub>16</sub> N <sub>2</sub> O <sub>6</sub> S	PVer_387.06_3.8	[M-H] <sup>-</sup>	387.0655	3.8	CID (15)	387.0662(35) 176.0356(16) 353.0784(100) 202.0149(25) 269.0570(11) 231.0414(12) 277.0297(15-9) 121.0296(21-0) 270.0646(16-2)
C <sub>18</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	PVer_341.16_20.1	[M+H] <sup>+</sup>	341.1604	20.1	CID (15)	341.1605(7) 324.1338(100) 296.1389(1)
C <sub>18</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	PVer_341.16_15.1	[M+H] <sup>+</sup>	341.1604	15.1	CID (15)	341.1601(12) 324.1337(100) 296.1388(1)
C <sub>18</sub> H <sub>22</sub> N <sub>4</sub> O <sub>3</sub>	PVer_343.17_12.7	[M+H] <sup>+</sup>	343.1747	12.7	CID (15)	343.1758(5) 326.1496(100)
C <sub>18</sub> H <sub>22</sub> N <sub>4</sub> O <sub>3</sub>	PVer_343.17_17.9	[M+H] <sup>+</sup>	343.1747	17.9	CID (15)	343.1758(3) 326.1495(100)
C <sub>18</sub> H <sub>24</sub> O <sub>7</sub>	PVer_351.14_18.1	[M-H] <sup>-</sup>	351.1448	18.1	CID (15)	351.1448(60) 307.1190(100) 249.0773(40) 333.1351(17)
C <sub>18</sub> H <sub>24</sub> O <sub>7</sub>	PVer_351.14_29.0	[M-H] <sup>-</sup>	351.1448	29.0	CID (15)	351.1450(25) 201.0924(20) 245.1188(100) 277.1449(15) 263.1293(6) 231.1033(21) 233.1188(2) 213.0923(1) 321.1344(6) 275.1291(5)
C <sub>18</sub> H <sub>24</sub> O <sub>7</sub>	PVer_351.14_24.6	[M-H] <sup>-</sup>	351.1448	24.6	CID (15)	351.1448(27) 321.1347(100) 307.1190(27) 245.0825(7) 249.0771(6) 319.1191(6) 205.0824(1)
C <sub>18</sub> H <sub>24</sub> O <sub>7</sub>	PVer_351.14_29.1	[M-H] <sup>-</sup>	351.1448	29.1	CID (15)	351.1448(20) 201.0925(25) 245.1189(100) 277.1448(16) 231.1031(19) 307.1551(16) 263.1291(6) 321.1355(6)
C <sub>18</sub> H <sub>24</sub> O <sub>7</sub>	PVer_351.14_24.1	[M-H] <sup>-</sup>	351.1448	24.1	CID (15)	351.1450(100) 321.1349(92) 249.0773(42) 307.1191(30) 245.0826(8) 319.1190(55) 205.0873(15) 333.1346(15) 231.0570(4)
C <sub>18</sub> H <sub>28</sub> O <sub>2</sub>	PVer_277.21_39.4	[M+H] <sup>+</sup>	277.2158	39.4	CID(25)	277.2164(20) 179.1431(100) 135.1168(42) 259.2060(30) 121.1011(15) 235.1696(6) 161.1325(11) 149.1324(27) 241.1954(20) 221.1538(15)
C <sub>18</sub> H <sub>32</sub> O <sub>4</sub>	PVer_335.22_36.2	[M+Na] <sup>+</sup>	335.2187	36.2	CID(20)	335.2187(100) 317.2096(30) 235.1309(22) 221.148(21) 251.1259(12)
C <sub>18</sub> H <sub>32</sub> O <sub>4</sub>	PVer_335.22_36.8	[M+Na] <sup>+</sup>	335.2187	36.8	CID(20)	335.2187(100) 317.2089(52) 235.1309(52) 221.1149(33) 243.2112(42) 261.2217(20)
C <sub>19</sub> H <sub>18</sub> N <sub>2</sub> O <sub>2</sub>	PVer_307.14_38.5	[M+H] <sup>+</sup>	307.1427	38.5	CID(25)	307.1439(22) 229.0971(100) 201.1023(4) 111.0554(4) 292.4467(1)
C <sub>19</sub> H <sub>18</sub> N <sub>2</sub> O <sub>3</sub> S	PVer_355.11_28.1	[M+H] <sup>+</sup>	355.1105	28.1	CID (15)	355.1135(4) 289.0967(50) 337.1000(100) 307.1074(85) 308.1150(2) 279.1125(12) 261.1019(1) 327.1156(10)

## Travail expérimental – Caractérisation du métabolome secondaire de *P. verrucosum*

C <sub>19</sub> H <sub>26</sub> O <sub>6</sub>	PVer_371.15_22.5	[M-2H+Na] <sup>-</sup>	371.1466	22.5	CID (15)	371.1466(45) 353.1375(100) 327.1579(15) 227.1057(6)
C <sub>19</sub> H <sub>37</sub> NO <sub>2</sub>	PVer_312.29_31.9	[M+H] <sup>+</sup>	312.2895	31.9	CID(25)	312.2895(12) 294.2792(100) 276.2684(4) 259.2417(4) 277.2524(5)
C <sub>20</sub> H <sub>18</sub> NO <sub>6</sub> Cl	PVer_404.09_33.1	[M+H] <sup>+</sup>	404.0897	33.1	CID (15)	404.0892(32) 358.0836(100) 239.0103(3) 341.0572(3) 386.0786(26)
	PVer_402.07_33.1	[M-H] <sup>-</sup>	402.0750		CID (15)	402.0756(65) 358.0855(100) 254.0229(4) 322.1090(4) 270.0696(4) 340.0682(4) 296.0801(3)
C <sub>20</sub> H <sub>19</sub> NO <sub>6</sub>	PVer_370.13_26.5	[M+H] <sup>+</sup>	370.1285	26.5	CID (15)	370.1286(4) 205.0492(16) 324.1226(100) 352.1172(32) 307.0959(4) 120.0805(4)
	PVer_368.11_26.5	[M-H] <sup>-</sup>	368.1130		CID (15)	368.1146(15) 324.1245(100) 220.0622(4) 249.0771(2) 306.1134(1) 350.1396(1)
C <sub>20</sub> H <sub>20</sub> N <sub>2</sub> O <sub>3</sub>	PVer_337.15_38.1	[M+H] <sup>+</sup>	337.1552	38.1	CID (25)	337.1552(3) 306.1360(100) 305.1281(20) 277.1331(10) 167.0851(5) 200.0707(13) 201.0742(10) 214.0862(4)
C <sub>20</sub> H <sub>20</sub> N <sub>2</sub> O <sub>4</sub>	PVer_353.15_32.0	[M+H] <sup>+</sup>	353.1480	32.0	CID (15)	353.1489(5) 305.1278(100) 336.1460(12) 118.0409(2) 275.1020(5) 322.1309(5) 250.1228(3)
C <sub>20</sub> H <sub>28</sub> O <sub>8</sub>	PVer_395.17_15.0	[M-H] <sup>-</sup>	395.1708	15.0	CID (15)	395.1708(100) 261.1866(12) 349.1661(65) 305.1762(25) 351.1815(20) 287.1656(3)
C <sub>20</sub> H <sub>30</sub> O <sub>5</sub>	PVer_349.20_22.5	[M-H] <sup>-</sup>	349.2012	22.5	CID (15)	349.2021(6) 305.1761(100) 287.1656(11)
C <sub>21</sub> H <sub>18</sub> N <sub>4</sub> O <sub>3</sub>	PVer_375.14_18.7	[M+H] <sup>+</sup>	375.1450	18.7	CID (25)	375.1449(3) 358.1183(100) 330.1231(1)
C <sub>21</sub> H <sub>18</sub> N <sub>4</sub> O <sub>3</sub>	PVer_373.13_17.0	[M-H] <sup>-</sup>	373.1295	17.0	CID (15)	373.1304(100) 302.0940(27) 355.1206(1)
C <sub>21</sub> H <sub>18</sub> N <sub>4</sub> O <sub>4</sub>	PVer_391.14_13.5	[M+H] <sup>+</sup>	391.1392	13.5	CID (25)	391.1398(5) 374.1130(100) 346.1182(2)
	PVer_389.12_13.5	[M-H] <sup>-</sup>	389.1258		CID (15)	389.1264(75) 318.0897(100) 278.0948(2) 1371.1153(5) 295.0843(5) 256.0732(2)
C <sub>21</sub> H <sub>18</sub> N <sub>4</sub> O <sub>4</sub>	PVer_391.14_21.0	[M+H] <sup>+</sup>	391.1392	21.0	CID (15)	391.1396(5) 374.1129(100) 346.1182(4) 149.0230(1)
C <sub>21</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	PVer_377.16_18.4	[M+H] <sup>+</sup>	377.1604	18.4	CID (25)	377.1603(4) 360.1336(100) 332.1392(1)
C <sub>21</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	PVer_377.16_12.8	[M+H] <sup>+</sup>	377.1604	12.8	CID (25)	377.1603(3) 360.1339(100) 332.1389(1)
C <sub>21</sub> H <sub>20</sub> N <sub>4</sub> O <sub>4</sub>	PVer_393.15_9.4	[M+H] <sup>+</sup>	393.1554	9.4	CID (15)	393.1551(30) 358.1181(12) 375.1448(100) 376.1284(22)
	PVer_391.14_9.4	[M-H] <sup>-</sup>	391.1400		CID (15)	391.1419(5) 373.1304(100) 300.0866(10) 285.0998(9) 256.0729(3) 320.1054(3) 348.1354(4) 263.0824(3) 154.0624(2)
C <sub>21</sub> H <sub>20</sub> N <sub>4</sub> O <sub>4</sub>	PVer_393.15_15.5	[M+H] <sup>+</sup>	393.1554	15.5	CID (15)	393.1541(15) 358.1181(2) 375.1447(100)
C <sub>21</sub> H <sub>21</sub> N <sub>3</sub> O <sub>3</sub>	PVer_364.16_3.6	[M+H] <sup>+</sup>	364.1656	3.6	CID (15)	364.1649(100) 308.1026(22) 200.1066(10) 346.1544(4) 252.1015(12)
C <sub>21</sub> H <sub>21</sub> N <sub>3</sub> O <sub>3</sub>	PVer_364.16_8.8	[M+H] <sup>+</sup>	364.1656	8.8	CID (15)	364.1650(25) 308.1027(100) 177.0654(14) 200.1067(7) 294.1232(2) 346.1543(4) 145.1280(2) 336.1540(2)
C <sub>21</sub> H <sub>30</sub> N <sub>4</sub> O <sub>4</sub>	PVer_401.22_25.8	[M-H] <sup>-</sup>	401.2185	25.8	CID (15)	401.2200(100) 262.1562(37) 172.1094(25) 245.1300(15) 228.1036(10) 353.0783(12) 203.0825(10)
C <sub>21</sub> H <sub>30</sub> N <sub>4</sub> O <sub>4</sub>	PVer_401.22_21.9	[M-H] <sup>-</sup>	401.2185	21.9	CID (15)	401.2200(100) 262.1562(28) 172.1094(19) 245.1301(12)
C <sub>22</sub> H <sub>27</sub> N <sub>3</sub>	PVer_334.23_15.8	[M+H] <sup>+</sup>	334.2271	15.8	CID (25)	334.2274(4) 203.1540(100) 217.1696(2)
C <sub>22</sub> H <sub>27</sub> N <sub>3</sub>	PVer_332.21_35.6	[M-H] <sup>-</sup>	332.2126	35.6	CID (15)	332.2126(100) 264.1148(13) 240.1504(5)
C <sub>23</sub> H <sub>21</sub> N <sub>5</sub> O <sub>3</sub>	PVer_416.17_14.6	[M+H] <sup>+</sup>	416.1712	14.6	CID (15)	416.1709(32) 399.1445(100) 130.0648(1) 270.0870(4) 287.1136(11)
C <sub>23</sub> H <sub>23</sub> N <sub>3</sub> O <sub>4</sub>	PVer_406.17_8.8	[M+H] <sup>+</sup>	406.1758	8.8	CID (15)	406.1753(32) 346.1544(100) 364.1649(40) 306.1232(16) 290.0921(2)
C <sub>24</sub> H <sub>28</sub> N <sub>4</sub>	PVer_373.24_16.0	[M+H] <sup>+</sup>	373.2385	16.0	CID (15)	373.2381(100) 242.1649(85) 111.0914(5) 256.1802(2)
C <sub>25</sub> H <sub>30</sub> N <sub>4</sub>	PVer_387.25_18.0	[M+H] <sup>+</sup>	387.2536	18.0	CID (25)	387.2542(82) 130.0648 (100) 271.1226(50) 257.1883(48) 330.1958(36) 201.1382(35) 156.0805(27) 342.1958(24) 258.1960(21) 187.1225(20)
C <sub>25</sub> H <sub>30</sub> N <sub>4</sub>	PVer_385.24_35.7	[M-H] <sup>-</sup>	385.2391	35.7	CID (15)	385.2404(67) 230.1668(100) 256.2821(55) 355.1625(17)
C <sub>25</sub> H <sub>30</sub> N <sub>4</sub>	PVer_385.24_37.2	[M-H] <sup>-</sup>	385.2391	37.2	CID (15)	385.2391(100) 230.1667(87) 256.1822(52) 156.0824(5) 341.1388(7) 328.1820(3)
C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	PVer_513.21_20.6	[M+Na] <sup>+</sup>	513.2119	20.6	CID (10)	513.2129(100) 496.1858(16)
C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	PVer_513.21_24.6	[M+Na] <sup>+</sup>	513.2119	24.6	CID (10)	513.2129(100) 496.1856(24)
C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	PVer_513.21_26.3	[M+Na] <sup>+</sup>	513.2119	26.3	CID (10)	513.2129(100) 496.1858(16)
C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	PVer_513.21_26.7	[M+Na] <sup>+</sup>	513.2119	26.7	CID (10)	513.2134(100) 496.1855(12)
C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	PVer_513.21_27.1	[M+Na] <sup>+</sup>	513.2119	27.1	CID (10)	513.2128(100) 496.1852(12)
C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	PVer_513.21_30.1	[M+Na] <sup>+</sup>	513.2119	30.1	CID (10)	513.2129(100) 496.1857(14) 351.1769(1)
C <sub>27</sub> H <sub>30</sub> N <sub>4</sub> O <sub>5</sub>	PVer_513.21_31.4	[M+Na] <sup>+</sup>	513.2119	31.4	CID (10)	513.2134(100) 496.1863(15)
C <sub>28</sub> H <sub>36</sub> N <sub>4</sub> O <sub>6</sub>	PVer_523.25_25.8	[M-H] <sup>-</sup>	523.2555	25.8	CID (15)	523.2562(100) 401.2204(5) 505.0936(2)
C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	PVer_511.29_18.0	[M+H] <sup>+</sup>	511.2909	18.0	CID (25)	511.2906(100) 265.1542(83) 219.1488(8) 247.1438(11) 394.2120(25) 412.2228(4) 349.1902(9) 493.2800(8) 295.1438(1)
	PVer_509.28_18.0	[M-H] <sup>-</sup>	509.2769		CID (15)	509.2769(100)
C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	PVer_511.29_18.7	[M+H] <sup>+</sup>	511.2909	18.7	CID (25)	511.2904(40) 265.1554(100) 219.1487(25) 247.1437(48) 346.2117(86) 493.2795(7) 166.0858(5) 394.2116(6) 301.1904(25)
	PVer_509.28_18.7	[M-H] <sup>-</sup>	509.2769		CID (15)	509.2769(100)
C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>6</sub>	PVer_525.27_13.0	[M-H] <sup>-</sup>	525.2720	13.0	CID (15)	525.2720(100)

## Travail expérimental – Caractérisation du métabolome secondaire de *P. verrucosum*

C <sub>30</sub> H <sub>38</sub> N <sub>4</sub> O <sub>2</sub>	PVer_509.29_26.2	[M+Na] <sup>+</sup>	509.2894	26.2	CID(20)	509.2899(3) 242.1651(100) 378.2171(45) 256.1805(5) 392.2327(2) 481.2791(1)
C <sub>30</sub> H <sub>38</sub> N <sub>4</sub> O <sub>2</sub>	PVer_509.29_25.2	[M+Na] <sup>+</sup>	509.2894	25.2	CID(20)	509.2895(2) 378.2171(100) 242.1651(40) 481.2797(10) 392.2320(5) 410.2061(5) 491.1917(2) 346.2126(2)
C <sub>30</sub> H <sub>38</sub> N <sub>4</sub> O <sub>2</sub>	PVer_509.29_23.0	[M+Na] <sup>+</sup>	509.2894	23.0	CID(20)	509.2895(3) 242.1347(100) 378.2168(45) 256.1786(6) 481.2792(4) 410.2070(4) 491.2005(4) 392.2334(4)
C <sub>30</sub> H <sub>38</sub> N <sub>4</sub> O <sub>2</sub>	PVer_509.29_24.4	[M+Na] <sup>+</sup>	509.2894	24.4	CID(20)	509.2898(2) 378.2159(100) 242.1648(86) 481.2793(13) 491.1976(7) 410.2065(7) 392.2333(6) 346.2109(4) 256.1801(1) 463.2033(1)
C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	PVer_548.29_17.4	[M-H] <sup>-</sup>	548.2892	17.4	CID (15)	548.2892(100)
C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	PVer_548.29_18.0	[M-H] <sup>-</sup>	548.2892	18.0	CID (15)	548.2892(100)
C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	PVer_548.29_19.1	[M-H] <sup>-</sup>	548.2892	19.1	CID (15)	548.2892(100)
C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	PVer_564.28_12.6	[M-H] <sup>-</sup>	564.2828	12.6	CID (15)	564.2828(100)
C <sub>31</sub> H <sub>20</sub> O <sub>11</sub>	PVer_567.09_26.3	[M-H] <sup>-</sup>	567.0941	26.3	CID (15)	567.0941(100)
C <sub>31</sub> H <sub>20</sub> O <sub>11</sub>	PVer_567.09_25.1	[M-H] <sup>-</sup>	567.0941	25.1	CID (15)	567.0941(100)
C <sub>31</sub> H <sub>40</sub> N <sub>4</sub> O <sub>2</sub>	PVer_523.30_27.7	[M+Na] <sup>+</sup>	523.3047	27.7	CID (15)	523.3038(3) 266.1177(100) 251.0939(15) 292.1331(5) 478.2491(5) 491.2156(4) 258.1965(4) 393.2407(3) 407.1753(3) 421.1904(3) 250.0864(3) 423.2073(3)
C <sub>31</sub> H <sub>40</sub> N <sub>4</sub> O <sub>2</sub>	PVer_523.30_27.4	[M+Na] <sup>+</sup>	523.3047	27.4	CID (15)	523.3058(3) 266.1177(100) 251.0942(12) 466.2490(6) 292.1332(6) 478.2487(3) 250.0863(3) 323.1750(2) 280.1332(2) 368.2335(2) 235.0997(2)
C <sub>32</sub> H <sub>36</sub> N <sub>4</sub> O <sub>2</sub>	PVer_507.28_38.5	[M-H] <sup>-</sup>	507.2760	38.5	CID (15)	507.2760(100)
C <sub>32</sub> H <sub>42</sub> O <sub>7</sub>	PVer_537.29_39.5	[M-H] <sup>-</sup>	537.2867	39.5	CID (15)	537.2867(100)
C <sub>32</sub> H <sub>42</sub> O <sub>8</sub>	PVer_553.28_38.7	[M-H] <sup>-</sup>	553.2814	38.7	CID (15)	553.2814(100)
C <sub>32</sub> H <sub>42</sub> O <sub>8</sub>	PVer_553.28_38.3	[M-H] <sup>-</sup>	553.2814	38.3	CID (15)	553.2814(100)
C <sub>37</sub> H <sub>42</sub> N <sub>4</sub> O <sub>5</sub>	PVer_623.32_36.3	[M+H] <sup>+</sup>	623.3215	36.3	CID (25)	623.3226(15) 373.2386(100) 242.1653(13)
	PVer_621.30_36.3	[M-H] <sup>-</sup>	621.3073		CID (15)	621.3073(100)
C <sub>38</sub> H <sub>42</sub> N <sub>4</sub> O <sub>10</sub>	PVer_713.28_33.4	[M-H] <sup>-</sup>	713.2847	33.4	CID (15)	713.2847(100)

*Table S1. Secondary metabolites detected from Penicillium verrucosum after culture on wheat grains and their MS/MS fragmentation in collision-induced dissociation (CID) mode. The fragmentation spectra are presented as a list of fragment ions with their relative intensity. <sup>a</sup> The chemical formulas were calculated after comparison of the m/z ratios obtained from the culture of P. verrucosum in <sup>12</sup>C wheat, 97% <sup>13</sup>C wheat and 53% <sup>13</sup>C / 96% <sup>15</sup>N wheat. <sup>b</sup> Secondary metabolites identifier (PVer\_m/z\_R<sub>T</sub>).*

## 2. OPTIMISATION DES RÉSEAUX MOLÉCULAIRES

---

### 2.1. Optimisation des données

Préalablement à la réalisation de réseaux moléculaires, les jeux de données ont été optimisés. Des MS/MS ciblées en mode CID et HCD ont été réalisées pour chaque métabolite secondaire observé chez *P. verrucosum*. Des premiers essais de réseaux moléculaires ont été réalisés à partir des données générées en DDA mais n'ont pas été concluants pour trois raisons. D'une part, ce mode n'a pas permis la fragmentation des métabolites minoritaires. D'autre part, pour certains composés, l'analyse en DDA procure un spectre MS/MS moyenné sur un faible nombre de scans et est donc de piètre qualité. Enfin, dans certains cas, les MS/MS automatiques ont été réalisées à des temps de rétention ne coïncidant pas avec l'intensité maximale du pic chromatographique et ne sont donc pas parfaitement représentatives de la molécule analysée.

Lors de la mise au point des méthodes d'analyse ciblées, les intervalles de temps au cours desquels chaque métabolite fongique a été fragmenté ont été précisément ajustés. De plus, les 98 analyses MS/MS ont été réparties sur trois injections distinctes afin de conserver un temps de scan suffisant et d'obtenir des spectres MS/MS moyens d'une dizaine de scans au minimum. Les isomères ont également été analysés lors d'injections distinctes dans l'objectif qu'ils soient différenciés lors de la génération de réseaux moléculaires. Enfin, les énergies de collision appliquées à chaque métabolite secondaire ont été optimisées pour que l'ion précurseur soit détecté avec une abondance relative minimale de 3% sur les spectres MS/MS, en visant une intensité relative idéale de 10%. Les réseaux présentés dans cette thèse ont été générés à partir des données optimisées en mode CID haute résolution.

### 2.2. Optimisation des paramètres du réseau moléculaire

Lors de la génération d'un réseau moléculaire, divers paramètres sont à optimiser. Leurs valeurs dépendent de la résolution des analyses et du niveau de tolérance souhaité pour établir une similarité entre deux spectres MS/MS. Les différents paramètres et leur définition sont présentés dans le Tableau 6.

MSCluster est un outil au cœur de la conception des réseaux moléculaires par le logiciel du GNPS. Développé par Frank *et al.*<sup>1</sup>, MScCluster est un outil destiné à faciliter le partage et la comparaison des données MS/MS. Dans le cadre de GNPS, MScCluster permet non seulement la comparaison des spectres MS/MS avec des données extérieures (bases de données, spectres de laboratoires étrangers), mais prend également part au traitement des données lors de la génération du réseau moléculaire lui-même. À cet effet, MScCluster est fondé sur une approche « d'archives spectrales » permettant de fusionner des spectres similaires. Ainsi, lors de la soumission des données brutes à GNPS, MScCluster fusionne tous les spectres issus du même ion précurseur en un unique spectre MS/MS moyen (à la tolérance du paramètre PIMT près) et présentant des niveaux de similarité supérieurs au paramètre score cosinus. Au-delà du regroupement des spectres MS/MS issus d'un même pic chromatographique dans une même analyse MS/MS ciblée, cette étape est utile pour la génération de réseaux à partir de plusieurs échantillons différents (exemple : échantillon test *versus* échantillon contrôle). En effet, elle permet de regrouper sous un même sommet (voir nomenclature des réseaux Figure 21) une molécule observée dans différents échantillons pour éviter la redondance d'information. En revanche, la fusion des spectres similaires issus d'ions précurseurs de même masse

exacte empêche la discrimination des différents isomères. Cette étape étant effectuée sur l'ensemble des fichiers soumis à l'outil du GNPS, cette limitation n'est pas contournée par l'analyse séparée des isomères dans différentes injections. Si son activation semble optionnelle, les nombreux essais réalisés à partir de nos données en l'absence de MSCluster ont systématiquement avorté.

PARAMÈTRES GNPS		Valeurs par défaut
Dépendants de l'analyse en spectrométrie de masse		
Tolérance de masse pour l'ion précurseur « Precursor ion mass tolerance (PIMT) »	Considérée lors de la comparaison des spectres et de la consultation des bases de données (en Da).	0,02
Tolérance de masse pour les ions fragment « Fragment Ion Mass Tolerance (FIMT) »	Considérée lors de la comparaison des spectres et de la consultation des bases de données (en Da).	0,02
Régissent la tolérance quant à la similarité des spectres MS/MS		
Score cosinus minimum « Min Pairs Cos »	Score cosinus minimum nécessaire à la liaison de deux spectres MS/MS.	0,7
Nombre minimum de fragments communs « Minimum Matched Fragment Ions (MMFI) »	Nombre minimum d'ions fragments que doivent partager deux spectres MS/MS pour être liés.	6
TopK des sommets « Network TopK »	Nombre maximum de voisins pour un sommet donné.	10
Taille minimum des clusters « Minimum Cluster Size »	Nombre minimum de spectres MS/MS dans un spectre MS/MS moyen pour que ce dernier soit considéré dans le réseau. Nécessite l'action de MSCluster.	2
Utiliser MSCluster « Run MSCluster »	Fusion des spectres MS/MS similaires avant génération du réseau.	Activé
Nombre maximum de sommets connectés « Maximum Connected Component Size »	Nombre maximum de sommets dans un graphe.	100
Filtration des ions fragment minoritaires « Filter stdDev Intensity »	Paramètre appliqué préalablement à MSCluster. Permet de calculer une intensité minimum des ions fragments en fonction des 25% moins intenses et supprime tous les fragments inférieurs à cette valeur. Action déconseillée par GNPS.	Désactivé
Intensité minimum des ions fragments « Minimum Peak Intensity »	Intensité minimum des ions fragments pour qu'ils soient pris en compte dans le calcul de similarité entre deux spectres MS/MS.	0
Filtration des pics à 50Da « Filter peaks in 50Da Window »	Élimine tous les pics qui ne sont pas parmi les 6 plus intenses dans une fenêtre de +/- 50 Da. Ne pas filtrer en cas d'analyse de petites molécules.	Activé

Tableau 6. Paramètres de génération de réseaux moléculaires par l'outil du GNPS.

Le score cosinus est le paramètre utilisé pour l'appréciation de la similarité entre deux vecteurs A et B en déterminant le cosinus de l'angle  $\theta$  entre eux :

$$\cos \theta = \frac{A \cdot B}{\|A\| \cdot \|B\|}$$

Lors de la formation des réseaux moléculaires, un score cosinus modifié, dont la formule n'est pas communiquée, intervient d'une part lors de la fusion des spectres MS/MS par MSCluster, et d'autre part lors de la comparaison des spectres MS/MS issus de  $m/z$  différents pour la génération des réseaux moléculaires. Un score proche de zéro indique une absence de similarité tandis qu'un score proche de 1 annonce une quasi superposition des deux spectres. Or, lors de l'optimisation des paramètres, il a été observé que certaines molécules identifiées de structures proches et présentant des spectres MS/MS très similaires étaient liées par un score cosinus faible. Ce biais peut être expliqué par le fait que le calcul du score cosinus prend en compte avec autant d'importance les intensités relatives des ions fragments et leurs rapports  $m/z$  lors de la comparaison de deux spectres MS/MS<sup>2</sup>. En effet, en spectrométrie de masse, les facteurs principaux permettant d'évaluer la similarité entre deux spectres sont la présence d'ions fragments communs ou les pertes de neutres communes. L'intensité des ions

résultants n'intervient qu'en second plan sachant que le simple ajout d'un groupement fonctionnel sur une structure peut considérablement déséquilibrer sa fragmentation et donc l'intensité de ses fragments.

Lors de l'optimisation des paramètres, des alternatives ont donc été recherchées pour s'affranchir de l'influence du score cosinus dans la génération des réseaux. L'abaissement du score cosinus n'a pas remis en cause la fusion des spectres par MSCluster puisque celle-ci est également contrôlée par la tolérance de masse de l'ion précurseur. En revanche, le réseau moléculaire résultant d'une analyse avec score cosinus minimum trop faible voyait tous ses sommets liés les uns aux autres et était difficilement interprétable. Le paramétrage de GNPS a été réalisé afin d'obtenir un réseau dans lequel les molécules de structures connues et de spectres MS/MS similaires étaient liées. Dans le cadre de l'analyse de *P. verrucosum*, les témoins positifs étaient l'ochratoxine A et l'ochratoxine B.

Les autres paramètres ont été optimisés de façon à trouver un compromis permettant l'analyse correcte de tous les types de spectres MS/MS soumis au logiciel. Les métabolomes secondaires fongiques comprennent des molécules de classes chimiques variées dans des concentrations très différentes. Cette hétérogénéité est à l'origine de spectres MS/MS variés, certains présentant une excellente sensibilité dans les conditions d'analyse mises en œuvre (Figure 34B), d'autres avec un faible ratio signal/bruit, mêlant les ions fragments à des signaux non spécifiques (Figure 34A). De plus, alors que certains ions fragmentent aisément à une faible énergie de collision, d'autres, plus stables, nécessitent des énergies de collision plus importantes. Enfin, alors que certains composés génèrent plusieurs ions fragments (Figure 34B), d'autres ne présentent qu'un petit nombre de fragmentations (Figure 34A).

Pour toutes ces raisons, l'optimisation du nombre minimum d'ions fragments communs nécessaires à l'établissement d'une similarité entre deux spectres MS/MS, ou encore l'intensité minimum des ions fragments à considérer doit être réalisée avec précaution.

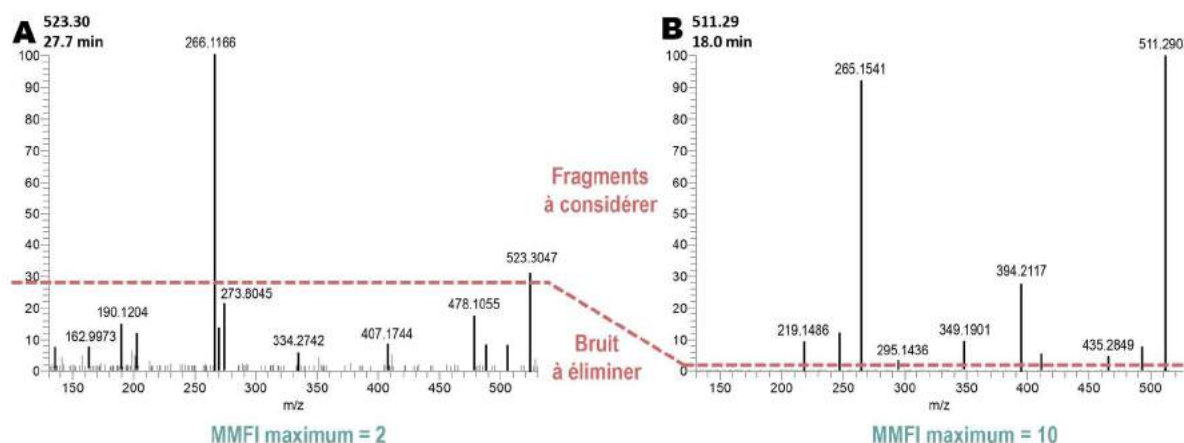


Figure 34. Optimisation des paramètres de GNPS en fonction de la diversité des spectres MS/MS à analyser. Exemple d'hétérogénéité entre les spectres MS/MS de (A) 523,047 à 27,7 min et de (B) 511,2903 à 18,0 min. La courbe pointillée délimite l'intensité du signal permettant de ne pas considérer le bruit de fond. Le Minimum Matched Fragment Ions (MMFI) maximum à paramétrer correspond au nombre de pics dépassant cette limite. Dans le cas d'un rapport signal/bruit faible, le paramétrage du MMFI peut empêcher la considération d'ions fragments de faible intensité. Par exemple, l'ion 407,1744 (A) sera ignoré en cas de paramétrage du MMFI à 2 alors qu'il se pourrait qu'il soit un ion fragment à considérer.

À l'issue de l'optimisation, les paramètres choisis pour la réalisation de réseaux moléculaires à partir des données CID haute résolution obtenues en electrospray positif chez *P. verrucosum* ont été les suivants : PIMT et FIMT = 0,0036 ; Min Pairs Cos = 0,01 ; Minimum Matched Fragment Ions = 1 ; TopK = 100 ; Minimum Cluster Size = 5 ; Run MScluster = activé ; Maximum Connected Component Size = 100 ; Filter stdDev Intensity = désactivé ; Minimum Peak Intensity = 0 ; Filter peaks in 50Da Window = désactivé.

### 2.3. Visualisation du réseau moléculaire

La visualisation des réseaux moléculaires a été réalisée avec le logiciel Cytoscape<sup>3</sup>. En raison de la fusion de spectres similaires par MScluster, les différents isomères (détectés à différents temps de rétention) n'ont pas été dissociés. Afin d'obtenir une cartographie globale du métabolome secondaire de *P. verrucosum*, les isomères sont ajoutés manuellement et liés au sommet présent par des arrêtes pointillées (voir la nomenclature des réseaux Figure 21). Cette opération est réalisée après comparaison visuelle des spectres MS/MS de ces isomères.

Lors de la visualisation du réseau moléculaire, nous avons souhaité ajouter l'intensité des ions associés à chaque sommet. En raison de la fusion des spectres similaires, les valeurs d'intensités des signaux relatifs aux différents sommets du réseau ne sont pas exploitables. Tout d'abord, elles correspondent aux intensités moyennées des isomères dont les spectres MS/MS ont été fusionnés par MScluster. De plus, les intensités sont extraites des profils MS/MS et non des profils chromatographiques issus du signal MS total (Total Ion Current (TIC)). Afin d'illustrer les différences d'intensité des ions sur le réseau moléculaire, les pics chromatographiques issus des TIC ont été intégrés manuellement et les aires (nommées A dans l'équation ci-dessous) ont été organisées selon une échelle logarithmique adaptée à la visualisation par Cytoscape.

$$\text{Taille}_{\text{sommet}} = 10 + 20 \times \log_{10}\left(\frac{A}{10^5}\right)$$

Ainsi, la taille de chaque sommet a été ajustée manuellement selon l'intensité de l'ion qu'il représente. Il est important de rappeler qu'en spectrométrie de masse, la différence d'intensité de deux ions n'est pas représentative de leur différence de concentration, chaque composé possédant un rendement d'ionisation particulier. De plus, un métabolite peut être soumis à un effet dit matrice dû à des composés présents au même moment dans la source d'ionisation (coélués). Le réseau moléculaire résultant est présenté dans l'Article 2 en fin de chapitre.

## 3. INTERPRÉTATION DES RÉSEAUX MOLÉCULAIRES

---

Comme décrit précédemment, l'élucidation de structures inconnues est facilitée lorsque leurs spectres MS/MS présentent des similarités avec le spectre MS/MS d'un composé connu. C'est pourquoi l'interprétation des réseaux moléculaires a débuté par l'étude des sommets adjacents aux sommets annotés. Afin de vérifier les potentielles similarités de structures entre l'ion 299,0913 et les ochratoxines par exemple, leurs spectres MS/MS sont comparés (Figure 35). Malgré leur liaison au sein d'un même graphe (voir nomenclature des réseaux Figure 21), l'absence d'ions fragments ou de pertes de neutres spécifiques ne permet pas d'appuyer l'hypothèse de structures communes entre les ochratoxines et l'ion inconnu.



## Travail expérimental – Caractérisation du métabolome secondaire de *P. verrucosum*

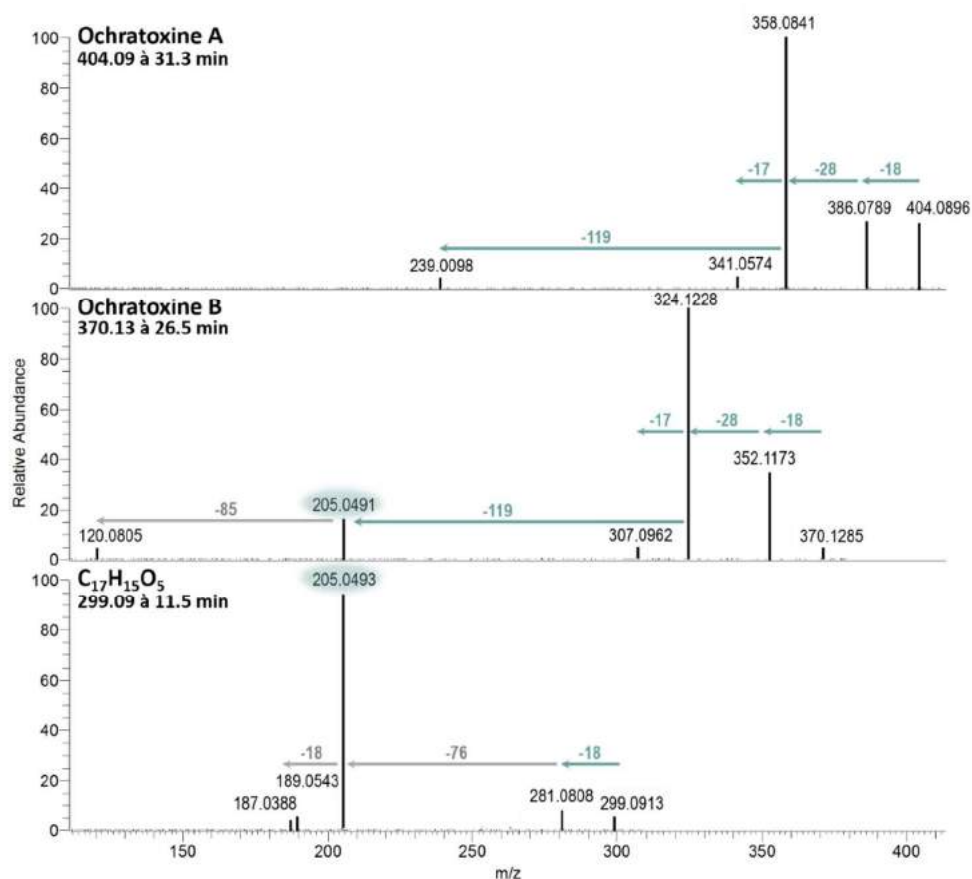


Figure 35. Comparaison des spectres MS/MS de l'ochratoxine A, de l'ochratoxine B et du métabolite C<sub>17</sub>H<sub>14</sub>O<sub>5</sub> détecté à 11,5 min.

La comparaison des spectres MS/MS des sommets annotés et des sommets adjacents inconnus est présentée dans le Tableau 7.

Annotation	m/z	T <sub>R</sub> (min)	Pertes communes	Fragments communs
OTB	370,1285	26,5		
?	299,0913	11,5	18	205,0491
Verrucine F	375,1450	18,7		
?	391,1392	13,4 et 20,9	17 ; 28	aucun
?	391,1392	13,4 et 20,9	17 ; 28	aucun
Verrucine A et/ou B	377,1604	18,4 et 12,8		
Verrucolone	147,0655	3,1		
?	129,0544	7,6	18	129,0544 ; 97,0281 ; 87,0438
Fungisporine A hydrolysée	511,2909	17,0		
?	511,2909	18,7	18 ; 46 ; 117 ; 246 ; 264 ; 292	511,2898 ; 493,2794 ; 465,2850 ; 394,2117 ; 265,1539 ; 247,1437 219,1484

Tableau 7. Analyse des similarités entre les sommets annotés et les sommets adjacents inconnus du réseau moléculaire de *P. verrucosum*.

#### 4. RÉFÉRENCES QUANT À L'ÉTUDE DE *P. verrucosum*

---

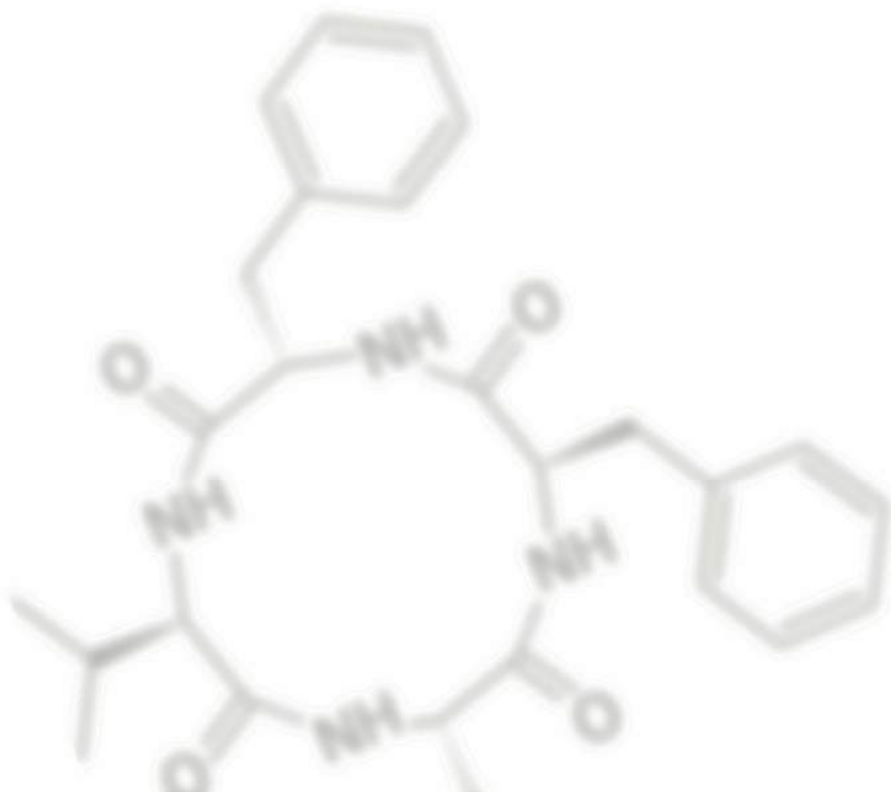
1. Frank AM, Monroe ME, Shah AR, et al. Spectral archives: extending spectral libraries to analyze both identified and unidentified spectra. *Nat Methods*. 2011;8(7):587-591.
2. Watrous J, Roach P, Alexandrov T, et al. Mass spectral molecular networking of living microbial colonies. *Proc Natl Acad Sci U S A*. 2012;109(26):E1743-52.
3. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-2504.



# 2

*Travail expérimental*

## **La production de fungisporines par *P. nordicum***





## 1. ARTICLE 3. RECHERCHE DE NOUVEAUX PEPTIDES CHEZ *P. nordicum*

---

Dans le cadre de cette thèse, la présence de fungisporines est recherchée chez les champignons producteurs d'ochratoxine A. En effet, les fungisporines pourraient être impliquées dans la toxicité et la virulence de ces champignons. De plus, la caractérisation de nouveaux métabolites secondaires pourrait permettre la découverte de nouveaux composés d'intérêts pharmaceutiques ou industriels. Dans le chapitre précédent, la fungisporine C ainsi que sept tétrapeptides linéaires potentiellement issus de l'hydrolyse de fungisporines ont été identifiés chez *P. verrucosum* (Article 2). La proximité phylogénique importante entre *P. verrucosum* et *P. nordicum*, ainsi que la détection chez *P. nordicum* d'un gène codant une protéine très similaire à HcpA (Article 3), nous permettent de suggérer que ce pathogène est également un producteur de fungisporines. L'objectif de ce chapitre consiste donc en l'analyse du métabolome secondaire de *P. verrucosum* pour la recherche de fungisporines. Les résultats obtenus permettront également la comparaison des métabolomes secondaires de *P. verrucosum* et de *P. nordicum* afin d'apprécier les similarités entre les métabolomes de deux moisissures très similaires sur les plans génétiques et morphologiques.

L'analyse du métabolome secondaire de *P. nordicum* a été réalisée dans les mêmes conditions que celles mises en œuvre lors de l'étude de *P. verrucosum* (Article 2). Outre l'aspect pratique de l'utilisation d'un substrat doublement marqué disponible au laboratoire, la culture des deux champignons sur un même milieu permet de comparer les données générées chez *P. nordicum* avec celles issues de *P. verrucosum*. Des réseaux moléculaires ont été générés à partir des métabolomes secondaires des deux espèces. Outre la visualisation des similitudes et différences entre ces derniers, les réseaux moléculaires ont permis la mise en évidence de nouveaux peptides potentiels. Leurs séquences ont alors été élucidées par spectrométrie de masse par interprétation des mécanismes théoriques de fragmentation peptidique présentés dans le chapitre Matériels et Méthodes. De plus, l'étude des fragmentations peptidiques par MS<sup>n</sup> a souligné un mécanisme de fragmentation non référencé auparavant. Les résultats de l'analyse du métabolome secondaire de *P. nordicum* figurent dans les Articles 3 et 4 en fin de chapitre.



## Characterization of 14 Fungisporins on Ochratoxin A Producing *Penicillium* species

Thaïs Hautbergue<sup>a,b</sup>, Olivier Puel<sup>a,\*</sup>, SouriaTadrist<sup>a</sup>, Lauriane Meneghetti<sup>a,b</sup>, Jean-Claude Tabet<sup>c,d</sup>, Laurent Debrauwer<sup>a,b</sup>, Isabelle P. Oswald<sup>a</sup>, Emilien L. Jamin<sup>a,b</sup>

<sup>a</sup> Toxalim (Research Centre in Food Toxicology), Université de Toulouse, INRA, ENVT, INP-Purpan, UPS, Toulouse, France

<sup>b</sup> Axiom platform, MetaToul-MetaboHUB, National Infrastructure for Metabolomics and Fluxomics, Toulouse, France

<sup>c</sup> CEA, iBiTec-S, SPI, LEMM, Metabohub Gif Sur Yvette, France

<sup>d</sup> Université Paris VI (UPMC), CNRS UMR 7201, Paris Cedex 05, France

\* Corresponding author: olivier.puel@toulouse.inra.fr

**ABSTRACT:** *Penicillium verrucosum* and *Penicillium nordicum* are the two ochratoxin A producing fungi infecting food and feed in temperate regions. Characterization of their secondary metabolomes is crucial to understand and control the toxic processes that led to great sanitary concerns and economic losses. In a previous work (Hautbergue *et al.* 2017), six fungisporins, including an unknown one, were characterized for the first time in *P. verrucosum*. The fungisporins are involved in fungal aerial growth and consequently in the contamination of foodstuffs by pathogenic fungi. Since these secondary metabolites could be development factors of *Penicillium* producing ochratoxin A, the synthesis of fungisporins by *P. nordicum* was evaluated in this study. For this purpose, an untargeted metabolomic approach based on a double stable isotope labeling was applied on *P. nordicum*, allowing the specific detection of its metabolites and the unambiguous determination of their chemical formulas. Analyses of the resulting extracts by High-Resolution Mass Spectrometry (HPLC-ESI-HRMS) showed that *P. nordicum* produces at least 92 secondary metabolites, 70 of which were totally unknown. Molecular networking of the secondary metabolomes of *P. verrucosum* and *P. nordicum* have led to the detection of 14 fungisporins never characterized from these fungi. Among them, five were unknown. Their structures were elucidated by mass spectrometry experiments (HPLC-ESI-MS<sup>n</sup>).

**Key words:** *Penicillium nordicum*; *Penicillium verrucosum*; mass spectrometry; stable isotope labeling; metabolomics; secondary metabolite; molecular network; fungisporin



### Introduction

Fungal secondary metabolomes consist of molecules produced by filamentous fungi providing them beneficial properties for adequate growth on their substrates. Over the past three decades, the need for new molecules of industrial concern is driving a renewed interest in exploration of fungal secondary metabolomes, particularly in pharmaceutical, cosmetic and alimentary industries. Moreover, about 400 fungal compounds are considered as mycotoxins [1]. Each year, the impact of these toxic compounds on global food security and economy is tremendous, particularly in a context of growing pesticides resistances and climate change [2]. Around 10% of the world crops are infested, leading each year to a destruction of food sources what could feed 600 million people *i.e.*, 8.5% of the world population [2]. Characterization of fungal secondary metabolomes is therefore crucial, especially since it has been assumed that the 15,600 fungal secondary metabolites currently identified represent less than 20% of the fungal secondary metabolomes [3].

Ochratoxin A is one of the eight mycotoxins internationally regulated [4]. This molecule is considered as the most potent renal carcinogen of natural origin [5] and has been classified as possible human carcinogen [6]. Ochratoxin A was related to urinary-tract tumors [7] and strongly suspected to be implicated in the fatal Balkan Endemic Nephropathy [8–10]. In Canada, annual economic losses resulting from wasting food and feed containing ochratoxin A at levels exceeding the security regulation limits amount to 260 million Canadian dollars [11]. In temperate regions, *Penicillium verrucosum* and *Penicillium nordicum* are the main ochratoxin A producing species [7]. While *P. verrucosum* is one of the most concerning saprophytic cereal pathogen fungi at these latitudes [12], *P. nordicum* is mostly referred as infecting dry-cured meat products, fish, cheese and occasionally wheat supplies [13]. Characterization of the secondary metabolites involved in the colonization of foodstuffs by these two fungi is crucial to understand and control their pathogenicity.

Untargeted metabolomics based on mass spectrometry (MS) stands as one of the most relevant analytical methods for the global characterization of fungal secondary metabolomes [14]. Analysis of fungal extracts by High-Resolution Mass Spectrometry (HRMS) allows the detection of both known and unknown metabolites even if present at trace level and contributes to their structural identification. However, in mass spectrometry analyses, secondary metabolites cannot be distinguished from contaminants. Moreover, the determination of an accurate mass by HRMS may lead to multiple possible chemical formulas, especially for higher mass metabolites [15]. Considering that the substrate is the only source of carbon and nitrogen atoms of a fungus, several approaches based on stable isotope labeling (SIL) have been developed to face these limitations [16]. Our laboratory developed a method based on the double isotope labeling of fungal secondary metabolites with both carbon 13 ( $^{13}\text{C}$ ) and nitrogen 15 ( $^{15}\text{N}$ ) by growing the fungus on three different substrates: (i) a non-labeled substrate (containing almost 100%  $^{12}\text{C}$ ), (ii) a fully labeled substrate (with almost 100%  $^{13}\text{C}$ ) and (iii) a doubly-labeled substrate (with 100%  $^{15}\text{N}$  and 50%  $^{13}\text{C}$ ). This strategy has the advantage to allow the specific detection of fungal secondary metabolites by the detection of their particular isotopic patterns and the unambiguous characterization of each chemical formula [17].

In our latest work, this method was applied for the characterization of the secondary metabolome of *P. verrucosum*. Among the 98 detected secondary metabolites, six fungisporins have been identified including a new one [18]. Fungisporins are fungal cyclic tetrapeptides whose first representative, the fungisporin A, was characterized by Studer *et al.* [19]. Ali *et al.* later demonstrated on *P. chrysogenum* that such compounds were involved in the aerial growth of the fungus [20]. Several of these

compounds as well as other derived of cyclic tetrapeptides were also characterized on *A. nidulans* [21]. In these studies, the fungisporins were detected in their original cyclic form and under several linear derivatives resulting from enzymatic ring opening. Recently, it was highlighted that the gene responsible of the synthesis of fungisporins was the most abundant biosynthetic gene cluster coding for nonribosomal peptide synthetases (NRPS) among 24 *Penicillium* species [22]. This observation has led to the hypothesis that such secondary metabolites are produced by most *Penicillium* species. To the best of our knowledge, this class of cyclic tetrapeptides as well as their linear derivatives have never been observed in ochratoxin A producing species. Considering the possible involvement of these molecules in the fungal growth, the identification of new fungisporins in ochratoxin A producing *Penicillium* may help the understanding of the colonization processes of foodstuffs by these fungi and the elaboration of strategies to control their pathogenicity.

### Material and Methods

#### 1.1. Chemicals and Reagents.

Formic acid as well as analytical solvents (acetonitrile, ethyl acetate and methanol) came from Fisher Scientific (Thermo Fisher Scientific, Illkirch, France). The solvents used for sample preparation were HPLC grade. Acetonitrile was MS grade. Ultra-pure water was generated from a Milli-Q system (Millipore, Saint Quentin en Yvelines, France). Ochratoxin A was purchased from Sigma (Sigma-Aldrich, Saint-Quentin Fallavier, France) and ochratoxin B was purchased from Fermentek (Fermentek Ltd, Jerusalem, Israel). The peptides were ordered from GeneCust (GeneCust, Dudelange, Luxembourg). All standards were dissolved in methanol for qualitative analysis by HPLC-MS.

#### 2.2. Characterization of *Penicillium nordicum*.

The cultures of *P. nordicum* were performed from the strain NRRL 6062. The identity of the strain was firstly verified by morphological characterization according to the monograph of *Penicillium* genus [13]. Secondly, the genomic characterization of the strain was performed. For this, the DNA of the fungus was extracted after 5 days of culture on YES medium (2.5 g of yeast extract, 10 g of glucose and 500 ml of distilled water). The mycelium was crushed and incubated in an extraction buffer (CTAB 1%, Tris-HCl pH 8 100 mM, EDTA pH 8 20 mM and NaCl 1.4 M) and chloroform, phenol, and isopropanol were added. The pellet was washed with ethanol 70% and suspended in water. Amplification was carried out with ITS (ITS 4 and 5) and  $\beta$ -tub (a and b) primers following the recommendations of Adjovi *et al.* [23]. Gel electrophoresis was performed from the PCR result to isolate each fragment that was then purified with GenElute PCR Clean-Up Kit (Sigma-Aldrich). Nanodrop (ND-1000 software, Thermo Scientific) was used to measure the DNA concentration. DNA sequences were compared by sequences alignment. Finally, the identification of *P. nordicum* was confirmed by the identification of ochratoxin A and anacine in its secondary metabolome (see results). These three criteria allowed the certification of the identification of *P. nordicum* according to the definition of Larsen *et al.* [24].

#### 2.3. Production of $^{13}\text{C}$ - and $^{13}\text{C}/^{15}\text{N}$ Labeled Wheats

The wheat variety Caphorn of *Triticum aestivum* was uniformly enriched with stable isotopes according to a previously developed method [25]. During 6 months, the wheat was grown in hermetically sealed chambers with  $^{13}\text{C}$ -labeled carbon dioxide  $^{13}\text{CO}_2$  with or without  $^{15}\text{N}$ -labeled nutrients ( $^{15}\text{NO}_3^-$  and  $^{15}\text{NH}_4^+$  salts) to generate two different enrichments. The first wheat (referred as

<sup>13</sup>C wheat) was grown with only <sup>13</sup>CO<sub>2</sub> to be enriched with nearly 100% of <sup>13</sup>C. The second (referred as <sup>13</sup>C/<sup>15</sup>N wheat) was grown with 50% of <sup>13</sup>CO<sub>2</sub> and 50% of <sup>12</sup>CO<sub>2</sub> as source of carbon atoms and with 100% of <sup>15</sup>N-labeled nutrients as source of nitrogen atoms. A Delta V Advantage isotope ratio mass spectrometer (Thermo Fisher Scientific) was used to measure the isotopic enrichment: 96.80% ± 0.15% for the <sup>13</sup>C wheat, 53.40% ± 0.3% <sup>13</sup>C and 96.80% ± 0.07% <sup>15</sup>N for the <sup>13</sup>C/<sup>15</sup>N wheat.

### 2.4. Cultures on wheat grains

Natural and labeled wheat grains were soaked in water and autoclaved. Before inoculation, an HC2-AW measurement device (Rotronik AG, Basserdorf, Switzerland) was used to measure the water activity (*a<sub>w</sub>*) of the grains. Thirty grams of each type of grains were then placed in sterile Petri dishes (Sarstedt, Marnay, France) in thin layer and the water activity was adjusted with sterile water to reach an *a<sub>w</sub>* of 0.98. The first dish contained <sup>13</sup>C wheat, the second dish received <sup>13</sup>C/<sup>15</sup>N wheat and two other dishes were filled with <sup>12</sup>C wheat. Cultures of the fungus were performed in the first three plates by an inoculation with 0.250 mL of a spore solution (at 10<sup>5</sup> spore/mL). The second <sup>12</sup>C wheat Petri plate was not inoculated but was stored in the same conditions as the other Petri dishes to constitute a control. The fungal cultures on natural and labeled wheats as well as the control were stored in the dark without shaking for one month at 25°C.

### 2.5. Metabolites Extraction

The fungal cultures and the control were separately extracted with ethyl acetate by shaking on a shaking table IKA Labor technik HS501 (IKA, Staufen, Germany) at 180 rpm during 36 h. Then, a filtration step allowed the elimination of the mycelium and the wheat grains or the meat. Crude extracts were evaporated and the resulting residues were resuspended in 400 µL of methanol. Before analysis by chromatography coupled with mass spectrometry, the four samples were filtered with 0.45 µm Whatman filters (Whatman, GE Healthcare, Kent, U.K.).

### 2.6. Liquid Chromatography, Mass Spectrometry Analyses and Ultra-Violet Analyses

The samples were analyzed by coupling HPLC with HRMS. The chromatographic separation was performed with an Ultimate 3000 system (Thermo Fisher Scientific, Les Ulis, France). The phase A of the mobile phase contained 0.1% of formic acid in water and the phase B contained 100% acetonitrile. The two solvents were eluted at the flow rate of 0.2 mL/min according to the following gradient: 0 min 20% of B, 30 min 50% of B, from 35 to 45 min 90% of B, from 50 to 60 20% of B. A volume of 10 µL of the injected sample was separated with a Luna C18 column (150 mm × 2.0 mm, 5 µm) (Phenomenex, Torrance, CA, U.S.A.). HRMS analyses were performed with a LTQ Orbitrap XL hybrid high-resolution mass spectrometer (Thermo Fisher Scientific, Les Ulis, France) in positive and negative electrospray ionization (ESI) modes. Ionization parameters were set as follows: spray voltage +4.5 kV, capillary temperature 350 °C, sheath gas (N<sub>2</sub>) flow rate 40 a.u. (arbitrary units), auxiliary gas (N<sub>2</sub>) flow rate 6 a.u. for the positive mode, and spray voltage -3.7 kV, capillary temperature 350 °C, sheath gas flow rate 30 a.u. (N<sub>2</sub>), auxiliary gas flow rate 10 a.u. (N<sub>2</sub>) for the negative mode. High-resolution mass spectra were acquired at a resolution of 60,000 from *m/z* 50 to *m/z* 800. UV Analysis of the fungal extract was also performed after a chromatographic separation on an Ultimate 3000 system (Thermo Fisher Scientific, Les Ulis, France). The high-resolution mass analyzer was calibrated in each ionization mode based on the supplier's protocol and calibration mixtures (Thermo Fisher Scientific, Les Ulis, France).

During the above-described analysis, MS/MS experiments were performed in data-dependent acquisition mode (DDA) in low resolution (ITMS).

After data analysis, targeted high-resolution MS/MS experiments were performed in both resonant and non-resonant modes on each secondary metabolites. In this article, CID refers to MS fragmentations in resonant processes and HCD refers to non-resonant fragmentations according to the nomenclature used by ThermoQuest. The HCD mode was used due to the absence of low mass cut-off and pseudopotential well. HCD and CID modes provide complementary structural information. Whereas CID mode lead to competitive fragmentations, HCD mode lead to consecutive fragmentation. [Parameters for MS/MS by resonant CID mode: resolution = 7500, MS<sup>n</sup> maximum ion time = 200 ms, isolation width = 1.5 *m/z*, Act.  $q_z = 0.250$  and activation time = 30 ms]. [Parameters for MS/MS by higher energy collision dissociation (HCD): resolution = 7500, last MS<sup>n</sup> step maximum ion time = 200 ms, isolation width = 1.5 *m/z* and activation time = 30 ms]. HCD and CID normalized collision energies (NCE) were adjusted to obtain appropriate fragmentation (10% of survival parent ion abundance) and maximum structural information.

Analyses by HPLC coupled with diode-array-detection (Ulti-mate 3000 system, Thermo Fisher Scientific, Les Ulis, France) at a wavelength of 220 nm were performed for the identification of some secondary metabolites. The chromatographic system and parameters were the same as for HPLC-HRMS analyses.

### 2.7. Identification of Secondary Metabolites and Molecular Networking

The home-made software MassToFormulaCompare (<http://www.toulouse.inra.fr/axiommm>) was used to compare the results from the natural and labeled wheats. The chemical formula of each secondary metabolite was determined with a measurement accuracy of 5 ppm and accepting the following elements: C, H, O, N, S, P, Cl, K and Na.

Identifications of metabolites were classified according to the Metabolomics Standard Initiative [26]. Metabolites identified at level 1 displayed the same retention time, chemical formula and product ion spectrum as the standard compound analyzed in the same conditions. The metabolites identified at level 2 displayed a product ion spectrum or an UV spectrum identical to this referenced in the literature. Some metabolites were identified at level 3 when they displayed product ion spectra coherent with their annotation. *In silico* fragmentations were performed using CFM-ID and compared with experimental product ion spectra to support the annotation of the metabolites identified at level 3 [27].

A molecular network generated with GNPS was created from the targeted high-resolution CID spectra of all secondary metabolites of *P. nordicum* detected in this study, together with the secondary metabolites detected in *P. verrucosum* in our later study [18]. The software parameters were set as follows: ion mass tolerance = 0.0036 Da, cosine score = 0.7, minimum match fragmentation = 1, maximum connected components = 100 and Top-K = 100, filter peaks removed. The resulting molecular network was visualized in Cytoscape 2.8.2 [28].

### 2.9. Analysis of cultures of *P. nordicum* on meat

Cultures of the strain NRRL 6062 were performed on cooked ham. The size of meat slices (0.5 cm thick) was adjusted to the diameter of petri dishes. Four ham slices were immersed in a solution of NaOCl 1% during 1 min, rinsed with sterile water twice and placed in separated petri dishes. Three meat slices were inoculated with 10  $\mu$ L of a spore solution (at 10<sup>6</sup> spore/mL) of *P. nordicum*. One meat

slice was not inoculated to constitute a control sample. The petri dishes were stored at 25°C in the dark for one week. The secondary metabolites were extracted and the samples were prepared as presented above. Targeted analyses were performed using a TSQ Vantage triple-stage quadrupole mass spectrometer (Thermo Fisher Scientific) using positive ESI mode based on two transitions for each secondary metabolite (including each time the major product ion and a specific product ion). Separation was performed on an Accela 600 LC system (Thermo Fisher Scientific, Les Ulis, France), using the same chromatographic parameters as above described. A volume of 10 µL of the extract was injected.

In a second time, the diffusion of the secondary metabolites in the substrate was investigated. For this, the strain NRRL 6062 was grown on cooked ham as previously described, but a cellophane sheet was placed on the meat slice before inoculation[29]. The fungus can take the nutrient through the cellophane. After one week of culture, the mycelium and the cellophane sheet were extracted separately from the meat slice. The secondary metabolites that can spread within the substrate were found in the slice meat and the secondary metabolites kept in the mycelium were found only in the extract of the cellophane/fungus part.

### 3. Results

#### 3.1. Detection of a fungisporin producing NRPS

Comparison of the amino acids sequences between the protein HcpA (CAP39139.1) involved in the synthesis of fungisporins in *P. chrysogenum* and all proteins referenced in GenBank was performed by a BLAST analysis. Fourteen *Penicillium* species presented a sequence with high homology (between 90% and 99%) and for an important recovery (between 99% and 100%), supporting the hypothesis of Nielsen *et al.* that many *Penicillium* species are able to produce fungisporins [22]. Particularly, the protein KOS44855.1 (GenBank) of *P. nordicum* displayed 91% of homology for a recovery of 100% with the protein HcpA. To ensure that these two proteins are involved in the synthesis of the same family of metabolites, NRPSpredictor2 was used to detect the adenylation domains of KOS44855.1 and compare them with the adenylation domains of HcpA [30]. The active sites of the adenylation domains of the two proteins displayed 97% homology (Table 1). Genetic investigation of *P. nordicum* suggested that this pathogenic fungus is able to produce an NRPS involved in the synthesis of tetrapeptides. Since genomic observations do not systematically reflect the expressed secondary metabolome in fungi, the secondary metabolome of *P. nordicum* and the production of tetrapeptides were investigated.

	Active Sites of Adenylation Domains	Location in the Sequence
<b>First Adenylation Domain of the NRPS</b>		
in HcpA of <i>P. chrysogenum</i>	GVHAFDASLAETCFVLAGEVNGYGPTEASVASVL	[401-534]
in KOS44855.1 of <i>P. nordicum</i>	GVHAFDASLAETCFVLAGEVNGYGPTEASVASVL	[216-349]
<b>Second Adenylation Domain of the NRPS</b>		
in HcpA of <i>P. chrysogenum</i>	RAFASDAFIVEASIALGGEQVLYGTCESSAASFH	[1943-2077]
in KOS44855.1 of <i>P. nordicum</i>	RAFASDAFIVEASIALGGEQVLYGTCESSAASFH	[1758-1892]
<b>Third Adenylation Domain of the NRPS</b>		
in HcpA of <i>P. chrysogenum</i>	TAMSFDVSI <b>L</b> EMSAITGGEINAYGPSEA <b>TVV</b> ATV	[2737-2870]
in KOS44855.1 of <i>P. nordicum</i>	TAMSFDVSI <b>M</b> EMSAITGGEINAYGPSEA <b>AVI</b> ATV	[2553-2686]
<b>Fourth Adenylation Domain of the NRPS</b>		
in HcpA of <i>P. chrysogenum</i>	SAFAFDAAVMEAFVLCGGEMNGYGPTETCVFAV <b>I</b>	[3843-3977]
in KOS44855.1 of <i>P. nordicum</i>	SAFAFDAAVMEAFVLCGGEMNGYGPTETCVFAV <b>V</b>	[3660-3794]

**Table 1.** Comparison of the active sites sequences of the adenylation domains detected by NRPSpredictor2 for HcpA in *P. chrysogenum* and KOS44855.1 in *P. nordicum*. Different amino acids are highlighted in bold.

### 3.2. HPLC-HRMS and MS/MS Analyses of the Three Wheat Cultures

The secondary metabolome of *P. nordicum* was characterized according to a protocol previously developed in our laboratory [17]. The fungus was grown on natural wheat,  $^{13}\text{C}$  wheat grains and  $^{13}\text{C}/^{15}\text{N}$  wheat grains during 30 days at  $25^\circ\text{C}$  in the dark and a control sample corresponding to non-infected natural wheat grains was carried out in the same conditions. The resulting secondary metabolomes were extracted and analyzed separately by HPLC-ESI-HRMS. To avoid overlapping of signals due to the complex isotopic composition of the labeled molecules, the extracts were not pooled before analysis. Then, the data mining of the three resulting fungal secondary metabolomes was carried out manually. The chromatograms and mass spectra resulting from the infected  $^{13}\text{C}/^{15}\text{N}$  labeled wheat were interpreted firstly to take advantage of the distinctive isotopic pattern showed by the ions labeled with 50% of  $^{13}\text{C}$  and 100% of  $^{15}\text{N}$ . The unlabeled contaminants generated during incubation or extraction were therefore rejected. Thus, 99 and 105 isotopic patterns were specifically detected in the mass spectra generated by the positive and negative electrospray ionization modes, respectively. Analysis of the chromatograms profiles showed that 15 (for the positive mode) and 25 (for the negative mode) of these specific signals were generated by adduct ions or in-source fragment ions. Thereafter, only protonated ions are considered [31]. Several sodium adduct ions were kept when no protonated ions were found. This led to a list of 84 features in the positive mode and 78 in the negative mode. The data mining was then conducted as described in our latest article by the analysis of the  $^{13}\text{C}$  sample and the  $^{13}\text{C}/^{15}\text{N}$  sample [18]. Finally, comparison of the analysis of  $^{12}\text{C}$  wheat infected by *P. nordicum* with the control sample allowed the rejection of plant metabolites (19 for the positive mode and 26 for the negative mode). Detection of 25 redundant signals between the positive and the negative ionization modes led to a final list of 92 secondary metabolites from *P. nordicum* (Table 2).

By knowing the accurate mass of a given molecule in the three different label conditions, it was possible to determine unambiguously its chemical formula. The number of carbon atoms contained in each metabolite was determined by subtracting the  $m/z$  ratio of  $^{12}\text{C}$  signal from these of the  $^{13}\text{C}$  signal. Comparison of  $^{12}\text{C}$ ,  $^{13}\text{C}$  and  $^{13}\text{C}/^{15}\text{N}$  accurate masses led to the calculation of the number of nitrogen atoms. Taking in account these two restrictions, and considering the seven golden rules [32], only one chemical formula remains common for the three accurate masses. In that way, the formulas of the 92 secondary metabolites of *P. nordicum* were determined (Table 2).

## Travail expérimental – La production de fungisporines par *P. nordicum*

proposed identification	molecular formula <sup>a</sup>	<sup>12</sup> C <i>m/z</i>	parent ion	R <sub>T</sub> (min)	proposed identification	molecular formula <sup>a</sup>	<sup>12</sup> C <i>m/z</i>	parent ion	R <sub>T</sub> (min)	proposed identification	molecular formula <sup>a</sup>	<sup>12</sup> C <i>m/z</i>	parent ion	R <sub>T</sub> (min)
Verrucolone (9)	C <sub>5</sub> H <sub>9</sub> NO	100.0755	[M+H] <sup>+</sup>	2.9		C <sub>18</sub> H <sub>22</sub> N <sub>4</sub> O <sub>4</sub>	381.1521	[M+Na] <sup>+</sup>	15.3	IFYV (14)	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>6</sub>	541.3039	[M+H] <sup>+</sup>	13.8
	C <sub>6</sub> H <sub>10</sub> O <sub>4</sub>	147.0649	[M+H] <sup>+</sup>	3.1		C <sub>18</sub> H <sub>24</sub> N <sub>2</sub> O <sub>2</sub>	301.1900	[M+H] <sup>+</sup>	37.7		C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>6</sub>	539.2860	[M-H] <sup>-</sup>	
PC-2 (6)	C <sub>10</sub> H <sub>13</sub> NO <sub>2</sub>	180.1016	[M+H] <sup>+</sup>	3.9	Ochratoxin A (1)	C <sub>19</sub> H <sub>37</sub> NO <sub>2</sub>	312.2896	[M+H] <sup>+</sup>	31.8	VYFI (10)	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>6</sub>	541.3039	[M+H] <sup>+</sup>	15.4
	C <sub>11</sub> H <sub>16</sub> O <sub>4</sub>	213.1120	[M+H] <sup>+</sup>	15.6		C <sub>20</sub> H <sub>18</sub> NO <sub>6</sub> Cl	404.0894	[M+H] <sup>+</sup>	33.2		C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>6</sub>	539.2860	[M-H] <sup>-</sup>	
LL-P880α (7)	C <sub>11</sub> H <sub>16</sub> O <sub>5</sub>	229.1071	[M+H] <sup>+</sup>	5.5	Ochratoxin B (2)	C <sub>20</sub> H <sub>19</sub> NO <sub>6</sub>	370.1283	[M+H] <sup>+</sup>	26.5	Fungisporin D (20)	C <sub>29</sub> H <sub>50</sub> N <sub>2</sub> O <sub>3</sub>	497.3741	[M+Na] <sup>+</sup>	43.6
	C <sub>11</sub> H <sub>16</sub> O <sub>5</sub>	229.1071	[M+H] <sup>+</sup>	6.1		C <sub>20</sub> H <sub>19</sub> NO <sub>6</sub>	368.1138	[M-H] <sup>-</sup>	39.4		C <sub>29</sub> H <sub>51</sub> NO <sub>2</sub>	446.3966	[M+H] <sup>+</sup>	
LL-P880β (8)	C <sub>11</sub> H <sub>18</sub> O <sub>4</sub>	215.1277	[M+H] <sup>+</sup>	13.9		C <sub>20</sub> H <sub>30</sub> O <sub>4</sub>	333.2062	[M-H] <sup>-</sup>	39.4		C <sub>30</sub> H <sub>37</sub> N <sub>5</sub> O <sub>4</sub>	532.2901	[M+H] <sup>+</sup>	36.4
	C <sub>11</sub> H <sub>18</sub> O <sub>5</sub>	231.1226	[M+H] <sup>+</sup>	5.7							C <sub>30</sub> H <sub>37</sub> N <sub>5</sub> O <sub>4</sub>	530.2750	[M-H] <sup>-</sup>	
	C <sub>12</sub> H <sub>9</sub> N <sub>3</sub> O <sub>2</sub>	226.0620	[M-H] <sup>-</sup>	15.2		C <sub>21</sub> H <sub>30</sub> N <sub>4</sub> O <sub>4</sub>	401.2183	[M-H] <sup>-</sup>	35.4	Cyclo(YWVV) (22)	C <sub>30</sub> H <sub>37</sub> N <sub>5</sub> O <sub>5</sub>	548.2845	[M+H] <sup>+</sup>	29.4
	C <sub>12</sub> H <sub>9</sub> N <sub>3</sub> O <sub>2</sub>	226.0620	[M-H] <sup>-</sup>	12.7		C <sub>21</sub> H <sub>30</sub> N <sub>4</sub> O <sub>4</sub>	401.2183	[M-H] <sup>-</sup>	25.8			C <sub>30</sub> H <sub>37</sub> N <sub>5</sub> O <sub>5</sub>	546.2707	
	C <sub>12</sub> H <sub>9</sub> N <sub>3</sub> O <sub>2</sub>	226.0620	[M-H] <sup>-</sup>	18.0		C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>2</sub>	345.2519	[M+H] <sup>+</sup>	27.0		C <sub>30</sub> H <sub>37</sub> N <sub>5</sub> O <sub>6</sub>	562.2649	[M-H] <sup>-</sup>	26.2
	C <sub>13</sub> H <sub>17</sub> NO	202.1236	[M-H] <sup>-</sup>	12.8		C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>2</sub>	345.2519	[M+H] <sup>+</sup>	23.5	VFWV (19)	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	550.3015	[M+H] <sup>+</sup>	17.4
	C <sub>14</sub> H <sub>26</sub> O <sub>5</sub>	275.1849	[M+H] <sup>+</sup>	18.1		C <sub>21</sub> H <sub>32</sub> N <sub>4</sub> O <sub>5</sub>	419.2286	[M-H] <sup>-</sup>	17.1			C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	548.2875	
	C <sub>15</sub> H <sub>24</sub> O <sub>2</sub>	237.1847	[M+H] <sup>+</sup>	17.2		C <sub>21</sub> H <sub>36</sub> O <sub>3</sub>	337.2723	[M+H] <sup>+</sup>	37.1	WVVF (17)	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	550.3015	[M+H] <sup>+</sup>	19.1
	C <sub>15</sub> H <sub>24</sub> O <sub>2</sub>	237.1847	[M+H] <sup>+</sup>	22.1		C <sub>22</sub> H <sub>28</sub> N <sub>4</sub> O <sub>5</sub>	449.1793	[M-2H+Na] <sup>-</sup>	25.4			C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	548.2875	
	C <sub>15</sub> H <sub>24</sub> O <sub>2</sub>	237.1847	[M+H] <sup>+</sup>	23.1		C <sub>23</sub> H <sub>26</sub> N <sub>2</sub> O <sub>4</sub>	415.1635	[M-2H+Na] <sup>-</sup>	17.9	VYVV (18)	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	566.2959	[M+H] <sup>+</sup>	12.3
	C <sub>15</sub> H <sub>24</sub> O <sub>3</sub>	253.1797	[M+H] <sup>+</sup>	9.7		C <sub>25</sub> H <sub>32</sub> N <sub>4</sub> O <sub>5</sub>	469.2409	[M+H] <sup>+</sup>	21.2			C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	564.2824	
	C <sub>15</sub> H <sub>24</sub> O <sub>3</sub>	253.1797	[M+H] <sup>+</sup>	22.1		C <sub>25</sub> H <sub>32</sub> N <sub>4</sub> O <sub>5</sub>	469.2409	[M+H] <sup>+</sup>	24.1	WVVY (16)	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	566.2959	[M+H] <sup>+</sup>	13.2
	C <sub>15</sub> H <sub>24</sub> O <sub>3</sub>	251.1650	[M-H] <sup>-</sup>	35.5		C <sub>25</sub> H <sub>40</sub> N <sub>4</sub> O <sub>5</sub>	477.3060	[M+H] <sup>+</sup>	16.9			C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	564.2824	
	C <sub>15</sub> H <sub>24</sub> O <sub>3</sub>	251.1650	[M-H] <sup>-</sup>	23.2		C <sub>26</sub> H <sub>42</sub> O <sub>7</sub>	467.2980	[M+H] <sup>+</sup>	30.5		C <sub>30</sub> H <sub>43</sub> N <sub>5</sub> O <sub>5</sub>	576.3165	[M+Na] <sup>+</sup>	21.0
	C <sub>15</sub> H <sub>24</sub> O <sub>3</sub>	253.1797	[M+H] <sup>+</sup>	23.2		C <sub>28</sub> H <sub>36</sub> N <sub>4</sub> O <sub>6</sub>	525.2682	[M+H] <sup>+</sup>	25.8		C <sub>30</sub> H <sub>43</sub> N <sub>5</sub> O <sub>5</sub>	574.3030	[M-2H+Na] <sup>-</sup>	
	C <sub>15</sub> H <sub>24</sub> O <sub>3</sub>	251.1650	[M-H] <sup>-</sup>	40.4			523.2543	[M-H] <sup>-</sup>			C <sub>30</sub> H <sub>43</sub> N <sub>5</sub> O <sub>6</sub>	592.3107	[M+Na] <sup>+</sup>	15.5
	C <sub>15</sub> H <sub>26</sub> O <sub>2</sub>	237.1857	[M-H] <sup>-</sup>	40.4			511.2906	[M+H] <sup>+</sup>	18.0		C <sub>30</sub> H <sub>43</sub> N <sub>5</sub> O <sub>6</sub>	578.2984	[M-H] <sup>-</sup>	
Sclerotigenin (5)	C <sub>16</sub> H <sub>11</sub> N <sub>3</sub> O <sub>2</sub>	278.0922	[M+H] <sup>+</sup>	9.9	VFFV (15)	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	509.2770	[M-H] <sup>-</sup>	18.0		C <sub>31</sub> H <sub>41</sub> N <sub>5</sub> O <sub>5</sub>	562.3030	[M-H] <sup>-</sup>	19.2
	C <sub>16</sub> H <sub>11</sub> N <sub>3</sub> O <sub>2</sub>	278.0922	[M+H] <sup>+</sup>	15.1			511.2906	[M+H] <sup>+</sup>	18.7		C <sub>31</sub> H <sub>41</sub> N <sub>5</sub> O <sub>5</sub>	562.3030	[M-H] <sup>-</sup>	
	C <sub>16</sub> H <sub>14</sub> O <sub>3</sub>	253.0862	[M-H] <sup>-</sup>	11.5	VVVF (21)	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	509.2770	[M-H] <sup>-</sup>	11.5		C <sub>31</sub> H <sub>42</sub> O <sub>6</sub>	531.2722	[M-H] <sup>-</sup>	37.3
	C <sub>17</sub> H <sub>10</sub> O <sub>5</sub>	295.0595	[M+H] <sup>+</sup>	11.2	VYFV (11)	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>6</sub>	527.2850	[M+H] <sup>+</sup>	12.7		C <sub>31</sub> H <sub>45</sub> N <sub>5</sub>	510.3554	[M+Na] <sup>+</sup>	36.9
	C <sub>17</sub> H <sub>10</sub> O <sub>5</sub>	293.0450	[M-H] <sup>-</sup>	11.3			525.2718	[M-H] <sup>-</sup>			C <sub>31</sub> H <sub>45</sub> N <sub>5</sub>	510.3554	[M+Na] <sup>+</sup>	
	C <sub>17</sub> H <sub>14</sub> N <sub>2</sub> O <sub>3</sub>	293.0923	[M-H] <sup>-</sup>	11.3		C <sub>28</sub> H <sub>42</sub> O <sub>5</sub>	481.2905	[M+Na] <sup>+</sup>	43.8		C <sub>32</sub> H <sub>40</sub> O <sub>6</sub>	521.2866	[M+H] <sup>+</sup>	39.5
	C <sub>17</sub> H <sub>26</sub> O <sub>4</sub>	315.1571	[M-2H+Na] <sup>-</sup>	38.6		C <sub>28</sub> H <sub>42</sub> O <sub>5</sub>	481.2905	[M+Na] <sup>+</sup>	44.2		C <sub>32</sub> H <sub>42</sub> O <sub>7</sub>	537.2838	[M-H] <sup>-</sup>	39.5
	C <sub>17</sub> H <sub>34</sub> O <sub>5</sub>	317.2320	[M-H] <sup>-</sup>	39.1			537.3050	[M+Na] <sup>+</sup>	24.4		C <sub>32</sub> H <sub>42</sub> O <sub>8</sub>	553.2810	[M-H] <sup>-</sup>	38.3
	C <sub>17</sub> H <sub>34</sub> O <sub>5</sub>	317.2320	[M-H] <sup>-</sup>	39.5		C <sub>28</sub> H <sub>42</sub> N <sub>4</sub> O <sub>5</sub>	535.2920	[M-2H+Na] <sup>-</sup>	24.4		C <sub>32</sub> H <sub>44</sub> O <sub>9</sub>	571.2899	[M-H] <sup>-</sup>	36.8
	C <sub>17</sub> H <sub>34</sub> O <sub>5</sub>	317.2320	[M-H] <sup>-</sup>	41.2			537.3050	[M+Na] <sup>+</sup>	23.7		C <sub>34</sub> H <sub>64</sub> O <sub>8</sub>	599.4507	[M-H] <sup>-</sup>	39.2
	C <sub>17</sub> H <sub>34</sub> O <sub>5</sub>	317.2320	[M-H] <sup>-</sup>	38.1		C <sub>28</sub> H <sub>42</sub> N <sub>4</sub> O <sub>5</sub>	535.2920	[M-2H+Na] <sup>-</sup>	23.1		C <sub>38</sub> H <sub>39</sub> N <sub>5</sub> O <sub>7</sub>	676.2778	[M-H] <sup>-</sup>	36.9
	C <sub>17</sub> H <sub>34</sub> O <sub>5</sub>	317.2320	[M-H] <sup>-</sup>	41.0			537.3050	[M+Na] <sup>+</sup>	23.1		C <sub>38</sub> H <sub>45</sub> N <sub>5</sub> O <sub>6</sub>	668.3415	[M+H] <sup>+</sup>	23.8
	C <sub>18</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	341.1606	[M+H] <sup>+</sup>	20.1		C <sub>28</sub> H <sub>42</sub> N <sub>4</sub> O <sub>5</sub>	535.2920	[M-2H+Na] <sup>-</sup>	20.1		C <sub>38</sub> H <sub>45</sub> N <sub>5</sub> O <sub>6</sub>	668.3415	[M+H] <sup>+</sup>	
Aurantiomide C (4)	C <sub>18</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	341.1606	[M+H] <sup>+</sup>	15.2			539.2860	[M+H] <sup>+</sup>	26.9					
	C <sub>18</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	339.1458	[M-H] <sup>-</sup>	15.2			537.2698	[M-H] <sup>-</sup>	26.9					
Anacine (3)	C <sub>18</sub> H <sub>22</sub> N <sub>4</sub> O <sub>3</sub>	343.1761	[M+H] <sup>+</sup>	12.7	IFFV (13)	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>5</sub>	525.3046	[M+H] <sup>+</sup>	19.5					
	C <sub>18</sub> H <sub>22</sub> N <sub>4</sub> O <sub>3</sub>	343.1761	[M+H] <sup>+</sup>	17.8	VFFI (12)	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>5</sub>	525.3046	[M+H] <sup>+</sup>	20.2					

**Table 2.** Secondary metabolites detected in *P. nordicum*, ranked in ascending number of carbon atoms. <sup>a</sup> The chemical formulas of molecules were determined by comparison of the *m/z* ratios obtained from cultures on natural wheat grains, on wheat grains labeled with 97% of <sup>13</sup>C and on wheat grains labeled with 53% of <sup>13</sup>C and 96% of <sup>15</sup>N. The number in brackets refers to the corresponding structures in Figure 1.

Based on a database of natural products (AntiBase (2012)[33]) and on the literature about *Penicillium* species, annotation of the chemical formulas led to the identification of some of the secondary metabolites. Ochratoxin A (compound 1 in Figure 1) and ochratoxin B (compound 2 in Figure 1), two well-known mycotoxins of *P. nordicum* [13,24], were identified at a level 1 by comparison with standard compounds. Since many natural products are not available as reference compounds [31], several metabolites were putatively annotated by comparison of their analytical results with literature. The chemical formulas  $C_{18}H_{22}N_4O_3$  ( $m/z$  343.1761),  $C_6H_{10}O_4$  ( $m/z$  147.0649) and  $C_{16}H_{11}N_3O_2$  ( $m/z$  278.0922) were identified as three metabolites known to be produced by *P. nordicum* [13], namely anacine (compound 3 in Figure 1), verrucolone (compound 9 in Figure 1) and sclerotigenin (compound 5 in Figure 1), respectively [34][35]. Identification of anacine was consolidated since the experimental UV and mass spectra were in agreement with the spectra reported in the literature [34]. The diastereoisomers of anacine were detected at two retention times (Table 2). The metabolite PC-2, known to be produced by *P. nordicum* [36], the other similar  $\alpha$ -pyrones LL-P880 $\alpha$  and LL-P880 $\beta$  (compound 6-8 in Figure 1) and the aurantiomide C (compound 4 in Figure 1) were identified at a level 3. As anacine, the diastereoisomers of aurantiomide C were detected at two retention times (Table 2).

In order to support the identifications at level 3, high-resolution positive CID spectra (Table S1) were compared with *in silico* spectra generated by CFM-ID. Although it should not be considered as proof of identification, this tool provides relevant supporting information. The *in silico* spectrum of LL-P880 $\alpha$  contained a major product ion at  $m/z$  197.1 which was in accordance with the major ion of the CID spectra of the compound detected at  $m/z$  215.1277 in *P. nordicum*. Three other fragment ions were common (at  $m/z$  183.1,  $m/z$  137.1 and  $m/z$  171.1). Similarly, four common fragment ions were found in the *in silico* spectrum of LL-P880 $\beta$  and seven common fragment ions were found in the *in silico* spectrum of PC-2, with the respective CID spectra. Similarly, two common fragment ions were found in the *in silico* spectrum of aurantiomide C (at  $m/z$  324.1337 and  $m/z$  296.1388) with the CID spectra of the two isomers of  $C_{18}H_{20}N_4O_3$  ( $m/z$  341.1606).

In accordance with the genetic observations, 13 fungisporins or their linear peptides derived from hydrolysis of fungisporins were annotated (compounds 10 to 22 in Figure 1). According to Klitgaard *et al.* [21] and Ali *et al.* [20] who characterized these compounds, some linear tetrapeptides may result from the enzymatic degradation of fungisporins. The cyclic forms of the Fungisporin D ( $m/z$  532.2901) and of the *cyclo*(YWVV) ( $m/z$  548.2845) were detected in the fungal extract (Supplementary data, Figure S1). Moreover, 11 linear peptides were annotated as hydrolyzed derived from fungisporins. The hydrolyzed fungisporin A (VFFV,  $m/z$  511.2906, compound 15 in Figure 1) was identified at a level 1 thanks to the analysis of a reference compound. The hydrolyzed *cyclo*(VFVF) (VFVF,  $m/z$  511.2906, compound 21 in Figure 1) was identified by comparison of its chemical properties with these reported from *P. verrucosum* in our previous study [18]. Other hydrolyzed fungisporins were identified at a level 1: the hydrolyzed fungisporin B (VYFV,  $m/z$  527.2850 compound 11 in Figure 1); the hydrolyzed *cyclo*(FFIV) (VFFI,  $m/z$  525.3046, compound 12 in Figure 1); the hydrolyzed *cyclo*(FFVI) (IFFV,  $m/z$  525.3046, compound 13 in Figure 1); the hydrolyzed *cyclo*(YFVI) (IYFV,  $m/z$  541.3039, compound 14 in Figure 1); the hydrolyzed *cyclo*(YFIV) (VYFI,  $m/z$  541.3039, compound 10 in Figure 1); two isomers of hydrolyzed fungisporin D (WVVF, compound 17 in Figure 1 and VFWV, compound 19 in Figure 1, both at  $m/z$  550.3015) and two isomers of hydrolyzed *cyclo*(YWVV) (VYWV, compound 18 in Figure 1 and WVVY, compound 16 in Figure 1, both at  $m/z$  566.2959). The stereochemistry of these compounds was determined by comparison with standard peptides. First, standard peptides corresponding to the above-mentioned sequences for which all the amino acids



were in levoratory configuration were purchase. The product ion spectra of these standards were similar to the product ion spectra of the fungisporins and their hydrolyzed metabolites, but their retention times were different. The standards peptides were therefore purchase with stereochemistry presented in Figure 1. This time, both the product ion spectra and the retention times were identical to the secondary metabolites produced by the fungus on wheat grains.

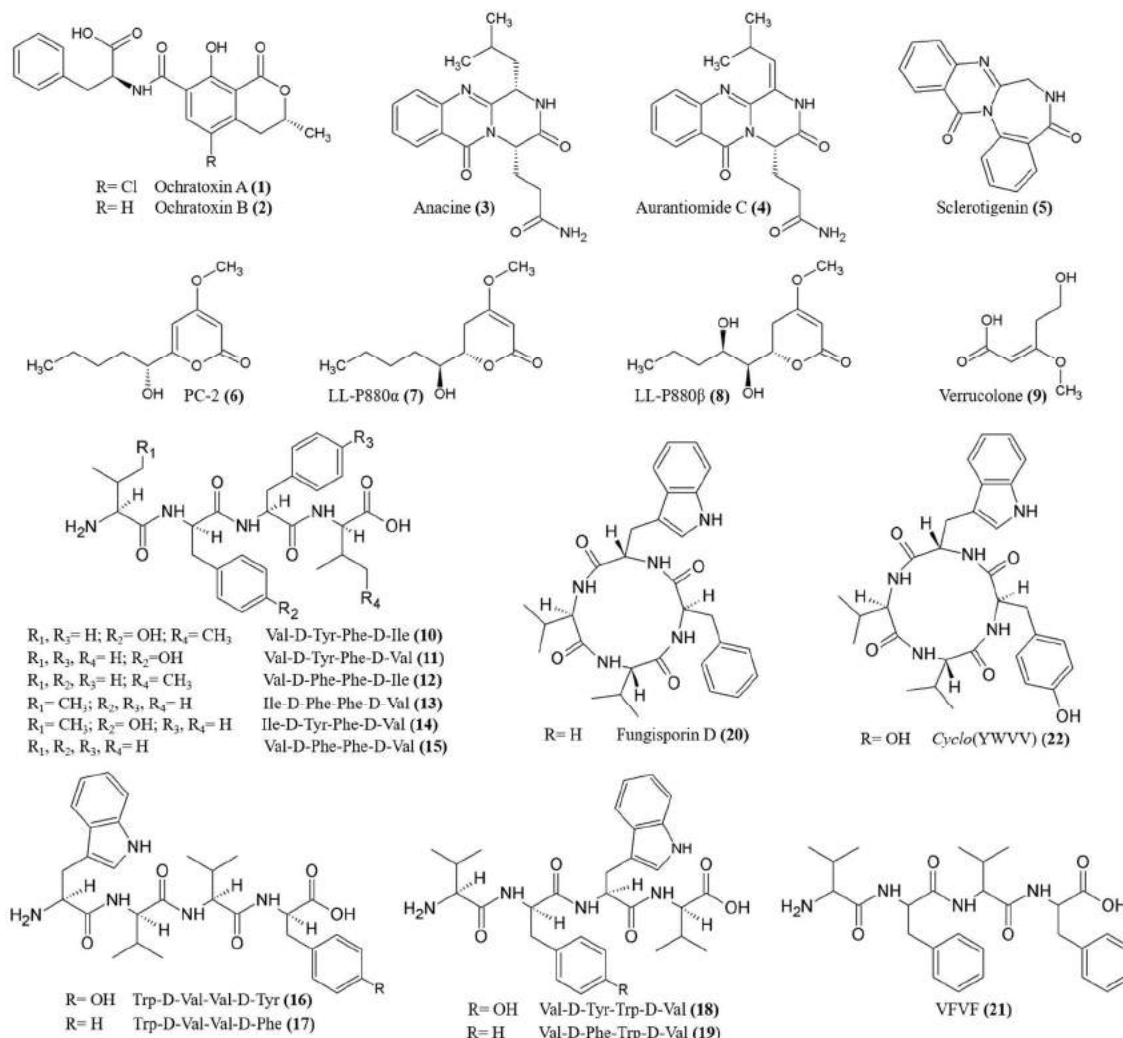


Fig. 1. Secondary metabolites identified from *P. nordicum*.

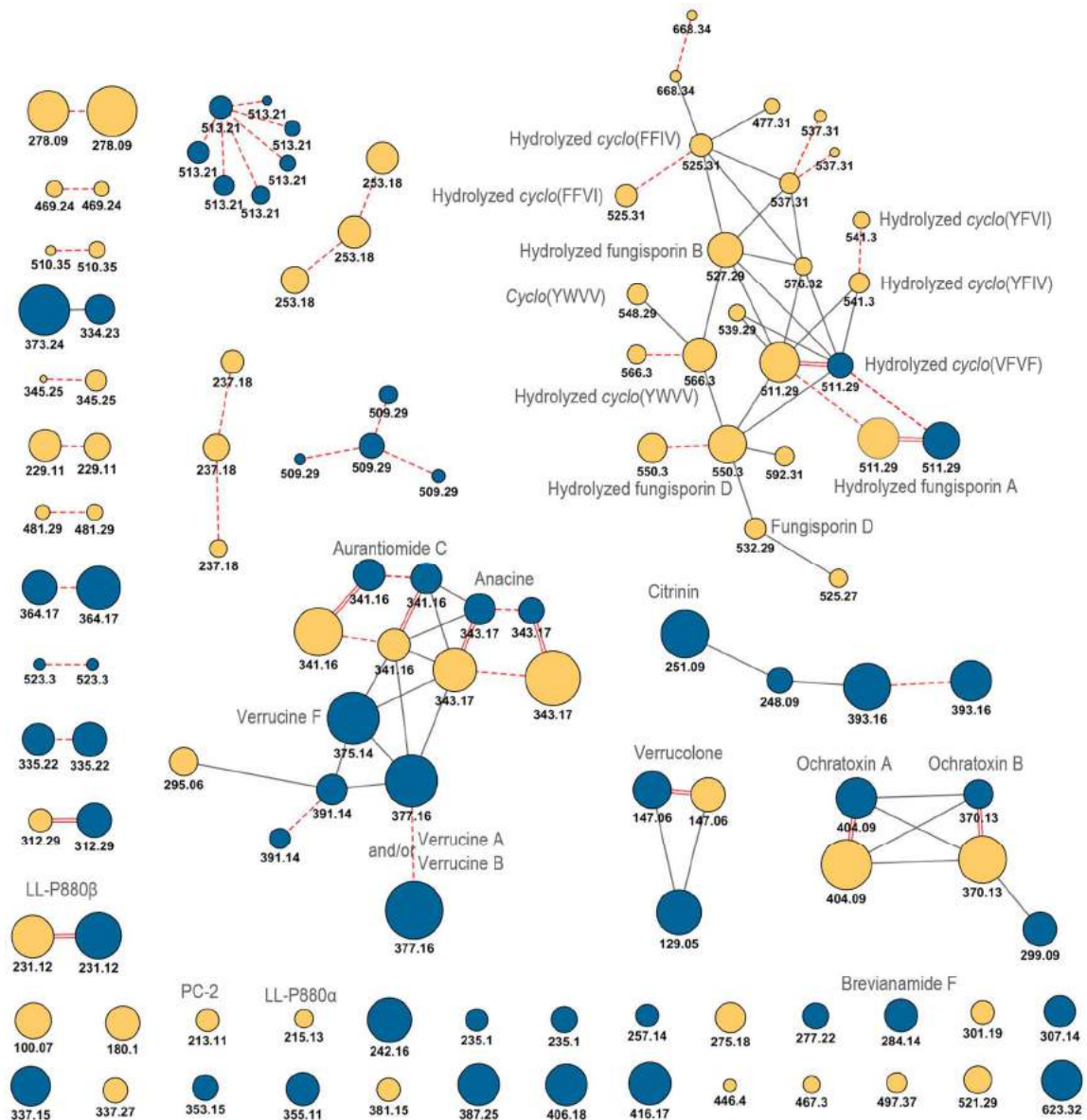
According to the assumption that such linear tetrapeptides may result from the hydrolysis of fungisporins, several isomers respecting a same amino acid alignment, such as VFWV and WVVF or VYVW and WVVY, were likely to come from a same cyclic tetrapeptide (fungisporin D and *cyclo*(YWVV) respectively). Therefore, 11 linear peptides likely generated by enzymatic hydrolysis of nine fungisporins were detected in *P. nordicum* (Table 3).

### 3.3. Molecular Networking for the Detection of New Tetrapeptides

The identification of 13 metabolites from the fungisporin family have raised the issue of the production of other unknown fungisporins by the fungus. One of the strategies for the detection of compounds belonging to a given chemical family could be performed by observation of their product ion spectra. Indeed, similar product ion spectra may come from similar molecules [37] and secondary metabolites fragmenting similarly may come from a same class of metabolites. Based on this

assumption, the GNPS Molecular Networking is a program dedicated to the treatment of metabolomic data for natural products research that provides a pioneering tool to organize hundreds of product ion spectra according to their similarity [38].

For the creation of such a molecular network, high-resolution targeted CID spectra were generated for each detected secondary metabolite of *P. nordicum* (Table S1) and processed in GNPS, together with the data from *P. verrucosum* generated in the same analytical conditions in our previous work [18]. The major difficulty for the molecular network construction came from the processing of datasets through a spectral archives approach, merging all the similar product ion spectra to create a single consensus spectrum before calculating their similarity score [39]. Therefore, this automatic clustering prevent the detection of isomers. To overcome this limit and to create a curated network, missing isomers were manually added in the network provided that their CID spectra were judged as very similar to the consensus spectra. To interpret the network (Figure 2) and to point out the production of potential new fungisporins by *P. nordicum*, the annotations proposed during the dereplication step were placed near to the corresponding nodes. Linking of similar structures, such as the ochratoxins or the quinazolines (verrucines, anacine and aurantiomide C), gave confidence in the GNPS parameters used to generate the molecular network. The identified fungisporins and their hydrolyzed derived were linked in a same graph composed of 25 nodes. The product ion spectra of the unknown metabolites linked with fungisporins were analyzed to elucidate their structures.

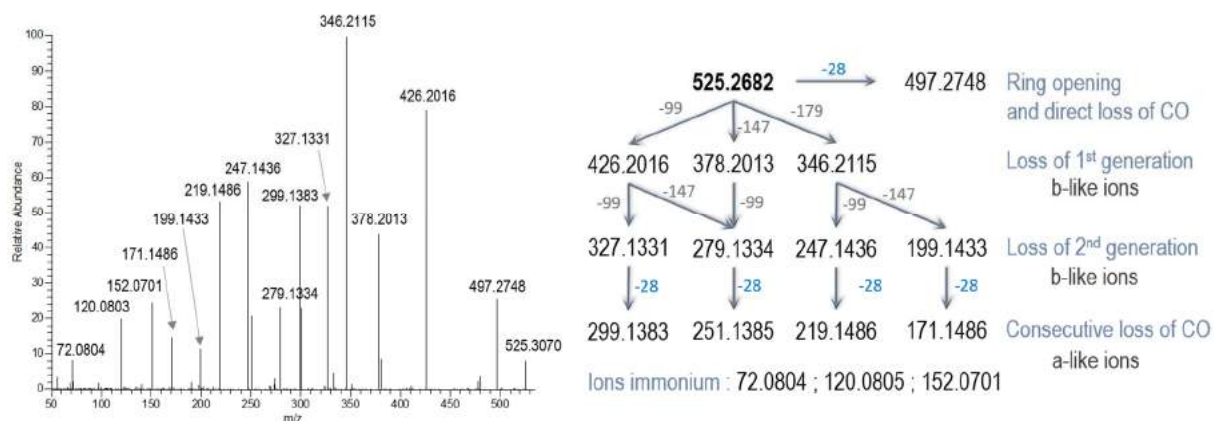


**Fig. 2.** Curated network obtained with targeted CID spectra performed on the metabolites detected in *P. nordicum* (yellow, light) and in *P. verrucosum* (blue, dark). Each node represents the CID spectrum of a secondary metabolite and is identified by the  $m/z$  of the precursor ion. Isomers manually added to overcome the spectral archive processing are linked by dotted line. Metabolites produced by the two fungi are linked by double lines. The nodes size (logarithmic scale) represents the intensity of the parent ions in the high-resolution mass spectra.

### 3.4. Characterization of a New Fungisporin

The metabolite  $C_{28}H_{36}N_4O_6$  detected at  $m/z$  525.2682 at 25.8 min was first annotated as the *cyclo*(YYVV), namely fungusporin C, but the analysis of its product ion spectra rejected this identification. The composition of this metabolites was determined according to the classical fragmentation mechanisms of cyclic peptides [40], [41] (Figure 3). The observation of immonium ions at  $m/z$  120.0803 and neutral losses of 147.0684 suggested the presence of a phenylalanine in the sequence and therefore rejected the annotation of fungusporin C. The presence of two residues of valine was supported by the fragment ions detected at  $m/z$  426.2016 and  $m/z$  327.1331 likely resulting from two consecutive losses of 99.0681. According to the hypothesis that this secondary metabolite is

a cyclic tetrapeptide, the fourth residue was a non-standard amino acid with a side chain composition of  $C_7H_7O_2$  that might correspond to a hydroxylated tyrosine. Since no losses of water was detected in the CID spectra, we suggested that the tyrosine was hydroxylated at the aromatic cycle and not at the benzyl position. The fragmentation mechanisms of cyclic peptides are much more complex than those of linear peptides since the ring opening may occur at each peptide bond [42]. Further sequential MS<sup>n</sup> experiments are therefore necessary to elucidate the sequence of this new fungisporin. However, the production of this compound in very low amount by *P. nordicum* did not allow the implementation of further MS<sup>n</sup> analyses in this study.



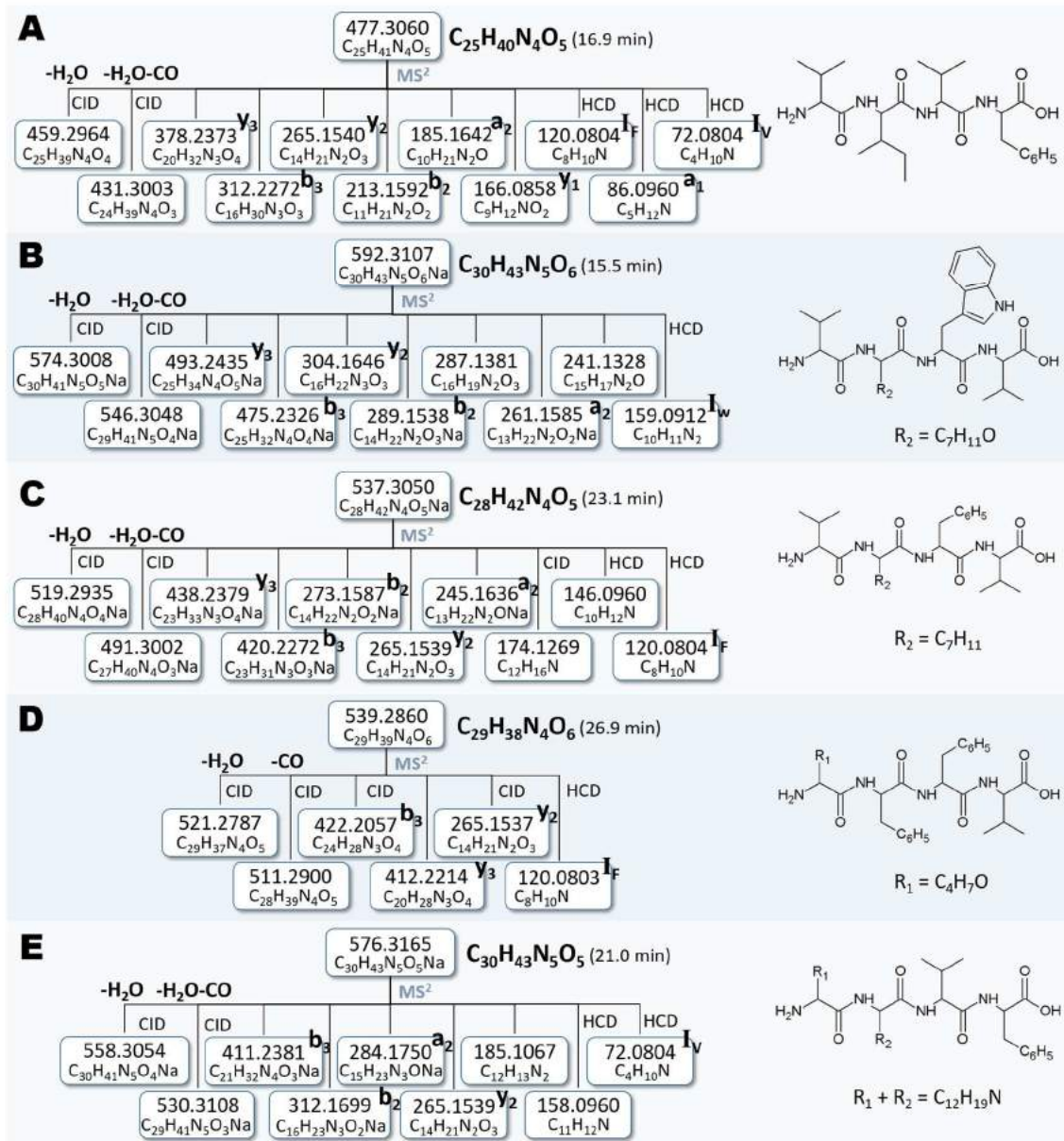
**Fig. 3.** Annotation of the HCD spectrum (NCE 20%) of the protonated metabolite  $C_{28}H_{36}N_4O_6$  detected at  $m/z$  525.2682 at 25.8 min.

The chromatographic profile and the molecular formula of this unknown metabolite detected in *P. nordicum* are identical to those of the ion annotated as fungisporin C in the previous study of *P. verrucosum* [18]. Here, we can therefore rectify the hypothesis about the secondary metabolome of the ochratoxin A producing *Penicillium* species and suggests that, although they don't produce fungisporin C, both *P. nordicum* and *P. verrucosum* produce the unknown cyclic tetrapeptide containing two valines, a phenylalanine and a non-standard amino-acid.

### 3.5. Detection of Five New Hydrolyzed Fungisporins and Two Pentapeptides

The structures of the unknown linear peptides highlighted by molecular networking were investigated based on classical peptidic fragmentation mechanisms [43][41]. Complementary information was provided by CID spectra and HCD spectra [41] (Table S1). The secondary metabolite detected at  $m/z$  477.3060 was characterized as the tetrapeptide VIVF (or VLVF) (Figure 4A). To support the manual *de novo* sequencing, the theoretical fragmentation of VIVF was simulated using the MS-Product tool of ProteinProspector [44]. Nine of the 11 fragment ions detected in the experimental product ion spectra were proposed by the software and only one theoretical fragment (corresponding to the  $a_3$  ion) was not found in the experimental product ion spectrum. To investigate the stereochemistry of the new metabolite detected at  $m/z$  477.3060, its analytical profile (MS/MS and retention time) was compared to these of several standards compounds: L-Val-L-Ile-L-Val-L-Phe, L-Val-L-Leu-L-Val-L-Phe, D-Val-L-Ile-D-Val-L-Phe, D-Val-L-Leu-D-Val-L-Phe, D-Val-L-Ile-L-Val-L-Phe and D-Val-L-Leu-L-Val-L-Phe. All these compounds displayed similar CID and HCD spectra but differed in their retention times. The protonated secondary metabolite at  $m/z$  477.3060 displayed the same product ion spectrum as the standards compounds, confirming the sequence of this new linear peptide.

However, its retention time was different from these of all the standard peptides. The stereochemistry of the amino acids could not be specified.



**Fig. 4.** Annotation of the high-resolution CID (NCE 15%) and HCD (NCE 20%) spectra of five unknown linear peptides detected in *P. nordicum* and proposition of corresponding structures. When non-specified, the product ions were detected in both CID and HCD modes. The product ions are annotated according to the nomenclature of Biemann [45].

Some of the unknown peptides displayed sequences with one or two non-standard amino-acids. The unknown peptides detected at  $m/z$  592.3107 and 537.3050 displayed similar structures. Interpretation of their product ion spectra indicated that they both consist of two valines (in C terminal and N terminal positions) and of an aromatic residue in third position (Figure 4B and C). The metabolite at  $m/z$  592.3107 displayed in second position an amino acid with a side chain  $C_7H_{11}O$  that could correspond to a tetrahydrotyrosine. Similarly, the metabolite at  $m/z$  537.3050 displayed an amino acid with a side chain that may correspond to a tetrahydrophenylalanine. This type of tetrahydroamino acids is known to be produced by microorganisms [46]. Two other isomers of the new peptide at  $m/z$  537.3050 were produced in very low amount. Their product ion spectra were uninterpretable since

they were polluted by the fragment ions of the major isomer described in this section. It is likely that the three isomers of  $C_{28}H_{42}N_4O_5$  resulted from different ring openings of a same cyclic tetrapeptide.

The other unknown peptide at  $m/z$  539.2860 displayed a N terminal residue with a side chain  $C_4H_7O$  that can correspond to a tetrahydrofuran group (Figure 4D). Finally, a new tetrapeptide was detected at  $m/z$  576.3165 and displayed two non-standard C terminal amino acids (Figure 4E).

The two unknown metabolites detected at  $m/z$  668.3415 were identified as new linear pentapeptides (Supplementary data, Figure S2). The metabolite detected at 24.2 min presented the sequence  $R_1GVVF$  with  $R_1$  as a non-standard amino acid with a side chain  $C_{15}H_{11}$ . The metabolite detected at 23.8 min was annotated as the peptide  $R_2GFFV$  with the side chain of  $R_2 = C_{11}H_{11}$ . The low amount of these two new compounds did not allow the analysis of their structure by  $MS^3$  or  $MS^4$ .

### 3.6. Target Detection of Fungisporines on Meat

The fungisporines and the linear peptides produced by *P. nordicum* on wheat grains were analyzed from cultures of the fungi on its natural substrate. In view of this, cooked ham slices were inoculated with the fungi. Since it was showed in *P. chrysogenum* that the fungisporins are produced during the first week of growing, the cultures of *P. nordicum* on meat were extracted after one week of growing. As a result, some of the known fungisporines were produced by *P. nordicum* on cooked ham slices in their cyclic or linear derived forms (Table 3). The new cyclotetrapeptide detected on wheat grains at  $m/z$  525.2682 was also produced by *P. nordicum* on cooked ham (Table 3). The new linear tetrapeptide at  $m/z$  477.3060 as well as the two pentapeptides at  $m/z$  668.3415 were detected on meat inoculated by *P. nordicum*.

To evaluate the diffusion of the peptides on the substrate, meat slices were coated with a cellophane sheet before being infected with *P. nordicum*. During one week, the fungus drew the nutrients from the meat through the cellophane sheet and produced secondary metabolites. Before extraction, the cellophane sheet containing the mycelium was removed. Separated analyses of the mycelium and of the meat slice were performed. As a results, the peptides were observed in both samples. This observation demonstrated that the peptides produced by *P. nordicum* may spread to the substrate and that the removing of the infected zone is not sufficient to protect the consumer from the ingestion of these compounds.

## 4. Conclusion

By the isotope labeling of the substrate of *P. nordicum*, 92 secondary metabolites were specifically detected and chemically characterized. Among these, 70 were totally unknown, supporting the assumption that almost 80% of the fungal secondary metabolites are still uncharacterized [3]. Chemical information about every detected secondary metabolite are presented in supplementary data (Table S1) to share the data for further *Penicillium* studies and to ensure the profitability of all these information: Fragmentation spectra are presented as a list of fragment ions with their relative abundance in HCD and/or CID modes.

The use of molecular networks has revealed the production of 25 similar compounds including 16 tetrapeptides of the family of fungisporins. For the first time, two fungisporines and 11 fungisporins derived were characterized on *P. nordicum* secondary metabolome (Table 3). Tandem mass

spectrometry experiments were performed to characterize the unknown molecules linked to fungisporines in the network. The linear tetrapeptide VIVF was identified as a new fungal secondary metabolite. Ali *et al.* reported that such fungal linear tetrapeptides might be generated by enzymatic hydrolysis of cyclic peptides. It was therefore suggested that VIVF illustrated the production of the new *cyclo(VIVF)* by *P. nordicum*. Since its sequence was not in accordance with the classical sequences of fungisporins *i.e.*, two consecutive aliphatic amino acids and two consecutive aromatic amino acids, it was expected that this new compound was produced by a distinct enzyme from a distinct gene. Five other tetrapeptides were partially characterized and contained one or two non-standards metabolites. The production of tetrapeptides by *P. nordicum* raises the question of the implication of these compounds in the contamination of food stocks by this pathogenic fungus. Indeed, studies have demonstrated that such linear or cyclic tetrapeptides may be involved in the development processes of *Penicillium* species by promoting the aerial growth of the fungus [20,22]. In order to control the environmental relevance of these secondary metabolites, the production of peptides by *P. nordicum* was investigated on meat. It was demonstrated that most of the known fungisporins as well as several of the new peptides were produced by *P. nordicum* when infecting its natural substrate. However, through the targeted research of these compounds, it was impossible to detect the secondary metabolites specifically produced by *P. nordicum* on meat, which requires a fully labelled meat.

Toxicological analyses are necessary to assess the potential pathogenicity of such secondary metabolites on human and animals. Depending on the results, targeted inhibitory strategies to inhibit the growth of the fungus as well as to control the amount of these tetrapeptides in the food may protect the population from the deleterious effect of *P. nordicum*. Such processes could also be applied to *P. verrucosum*, which was also demonstrated as a fungisporins producer. Finally, these secondary metabolites could present industrial interests since one of them, the linear FVVY likely originating from hydrolysis of fungisporin B, displayed antifungal activities similar to those of industrial fungicide benomyl [47]. Moreover, it was suggested by Ali *et al.* that fungisporins may display cardiac ion-channel blocking activities interesting for the treatment of hypertension and angina [20].

The generation of molecular networks, as well as the culture of the two fungi in identical conditions, allowed the comparison of the secondary metabolomes of *P. verrucosum* and *P. nordicum*. The phylogenetic proximity between the two species is so close that they were considered as only one species for a while [48]. However, only 11 secondary metabolites were found to be produced in common by the two fungi *i.e.*, less than 15% of their secondary metabolomes. This observation highlighted that two similar genomes may result in very different metabolomes, supporting the importance of the use of metabolic criterion for the characterization of fungal strains.

## Travail expérimental – La production de fungisporines par *P. nordicum*

Sequence	Molecular formula	m/z of molecular ion in ESI*	R <sub>T</sub> (min)	Common name of cyclic peptide	Production on wheat grains	Production on meat
<b>Known cyclic peptides detected</b>						
<i>Cyclo</i> (VFWV) or isomer <sup>b</sup> (20)	C <sub>30</sub> H <sub>37</sub> N <sub>5</sub> O <sub>4</sub>	532.2901	36.4	Fungisporin D	Yes	Not detected
<i>Cyclo</i> (YWVV) or isomer <sup>b</sup> (22)	C <sub>30</sub> H <sub>37</sub> N <sub>5</sub> O <sub>5</sub>	548.2845	29.4	<i>Cyclo</i> (YWVV)	Yes	Yes
<b>Known linear peptides detected</b>					<b>Potential corresponding cyclic peptide</b>	
VFFV <sup>a,b</sup> (15)	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	511.2906	18.0	Fungisporin A	Yes	Yes
VYFV <sup>a,b</sup> (11)	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>6</sub>	527.2850	12.7	Fungisporin B	Yes	Yes
VFWV <sup>a,b</sup> (19)	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	550.3015	17.4	Fungisporin D	Yes	Yes
WVVF <sup>a,b</sup> (17)	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	550.3015	19.1	Fungisporin D	Yes	Yes
VFVF <sup>a,b</sup> (21)	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	511.2906	18.7	<i>Cyclo</i> (VFVF)	Yes	Yes
IFFV <sup>b</sup> (13)	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>5</sub>	525.3046	19.5	<i>Cyclo</i> (FFVI)	Yes	Yes
VFFI <sup>b</sup> (12)	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>5</sub>	525.3046	20.5	<i>Cyclo</i> (FFIV)	Yes	Yes
VYFI <sup>b</sup> (10)	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>6</sub>	541.3039	15.4	<i>Cyclo</i> (YFIV)	Yes	Not detected
IYFV <sup>b</sup> (14)	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>6</sub>	541.3039	13.8	<i>Cyclo</i> (YFVI)	Yes	Yes
VYWV <sup>a,b</sup> (18)	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	566.2959	12.3	<i>Cyclo</i> (YWVV)	Yes	Yes
WVVI <sup>b</sup> (16)	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	566.2959	13.2	<i>Cyclo</i> (YWVV)	Yes	Not detected
<b>New cyclic peptides detected</b>						
<i>Cyclo</i> (R <sub>1</sub> VVF) <sup>a,b</sup> or sequence isomers R <sub>1</sub> with a side chain C <sub>7</sub> H <sub>7</sub> O <sub>2</sub>	C <sub>28</sub> H <sub>36</sub> N <sub>4</sub> O <sub>6</sub>	525.2682	25.8		Yes	Yes
<b>New linear peptides detected</b>					<b>Potential cyclic peptide of origin</b>	
VIVF (or VLVF) <sup>b</sup>	C <sub>25</sub> H <sub>40</sub> N <sub>4</sub> O <sub>5</sub>	477.3060	16.9	<i>Cyclo</i> (VIVF) or <i>cyclo</i> (VLVF)	Yes	Yes
VR <sub>2</sub> WV <sup>b</sup> R <sub>2</sub> with side chain C <sub>7</sub> H <sub>11</sub> O	C <sub>30</sub> H <sub>43</sub> N <sub>5</sub> O <sub>6</sub>	592.3107	15.5	<i>Cyclo</i> (VR <sub>2</sub> WV)	Yes	Not detected
VR <sub>2</sub> FV <sup>b</sup> R <sub>2</sub> with side chain C <sub>7</sub> H <sub>11</sub>	C <sub>28</sub> H <sub>42</sub> N <sub>4</sub> O <sub>5</sub>	537.3050	23.1	<i>Cyclo</i> (VR <sub>2</sub> FV)	Yes	Not detected
undetermined	C <sub>28</sub> H <sub>42</sub> N <sub>4</sub> O <sub>5</sub>	537.3050	23.7		Yes	Not detected
undetermined	C <sub>28</sub> H <sub>42</sub> N <sub>4</sub> O <sub>5</sub>	537.3050	24.4		Yes	Not detected
R <sub>1</sub> FFV <sup>b</sup> R <sub>1</sub> with side chain C <sub>4</sub> H <sub>7</sub> O	C <sub>29</sub> H <sub>38</sub> N <sub>4</sub> O <sub>6</sub>	539.2860	26.9	<i>Cyclo</i> (R <sub>1</sub> FFV)	Yes	Not detected
R <sub>1</sub> R <sub>2</sub> VF <sup>b</sup> Total of R <sub>1</sub> and R <sub>2</sub> side chains = C <sub>12</sub> H <sub>19</sub> N	C <sub>30</sub> H <sub>43</sub> N <sub>5</sub> O <sub>5</sub>	576.3165	21.0	<i>Cyclo</i> (R <sub>1</sub> R <sub>2</sub> VF)	Yes	Yes
R <sub>1</sub> GVVF <sup>b</sup> R <sub>1</sub> with side chain C <sub>15</sub> H <sub>11</sub>	C <sub>38</sub> H <sub>45</sub> N <sub>5</sub> O <sub>6</sub>	668.3415	24.2		Yes	Yes
R <sub>1</sub> GFFV <sup>b</sup> R <sub>1</sub> with side chain C <sub>11</sub> H <sub>11</sub>	C <sub>38</sub> H <sub>45</sub> N <sub>5</sub> O <sub>6</sub>	668.3415	23.8		Yes	Yes

**Table 3.** Fungisporins and other peptides produced by *P. verrucosum* (<sup>a</sup>) and *P. nordicum* (<sup>b</sup>). The number in brackets refers to the corresponding structures in Figure 1.

### Acknowledgements

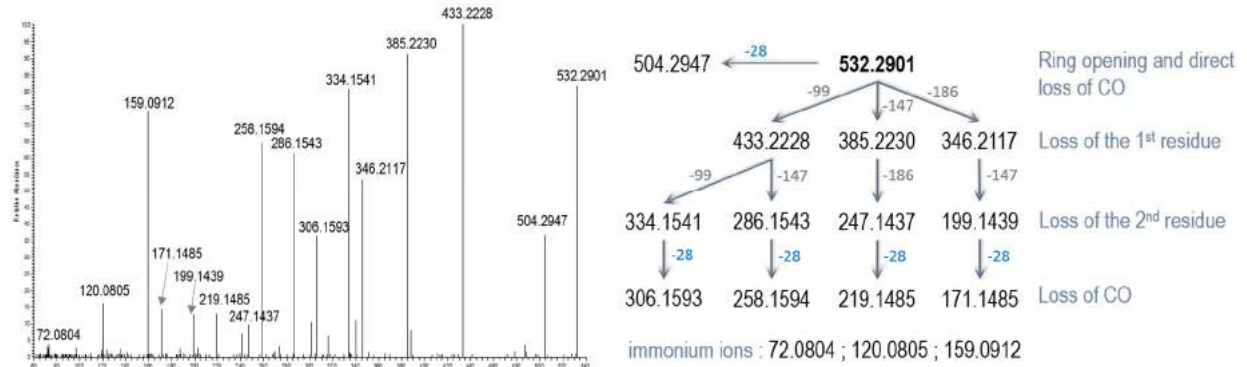
PhD doctoral fellowship was co-funded by INRA and French Minister of Higher Education and Research. Authors thank the Team Recherches Appliquées en Phytotechnologie, CEA, IBEB, Cadarache, Saint-Paul-les-Durance, France for labeled wheat production. The authors are grateful to Pascal Bourdaudhui who developed the software MassToFormulaCompare, and to Jean-Claude Tabet for his kind help for the mass spectrometry analyses and data monitoring. All MS experiments were performed on the instruments of the MetaToul-AXIOM platform, partner of the national infrastructure of metabolomics and fluxomics: MetaboHUB (MetaboHUB-ANR-11-INBS-0010).



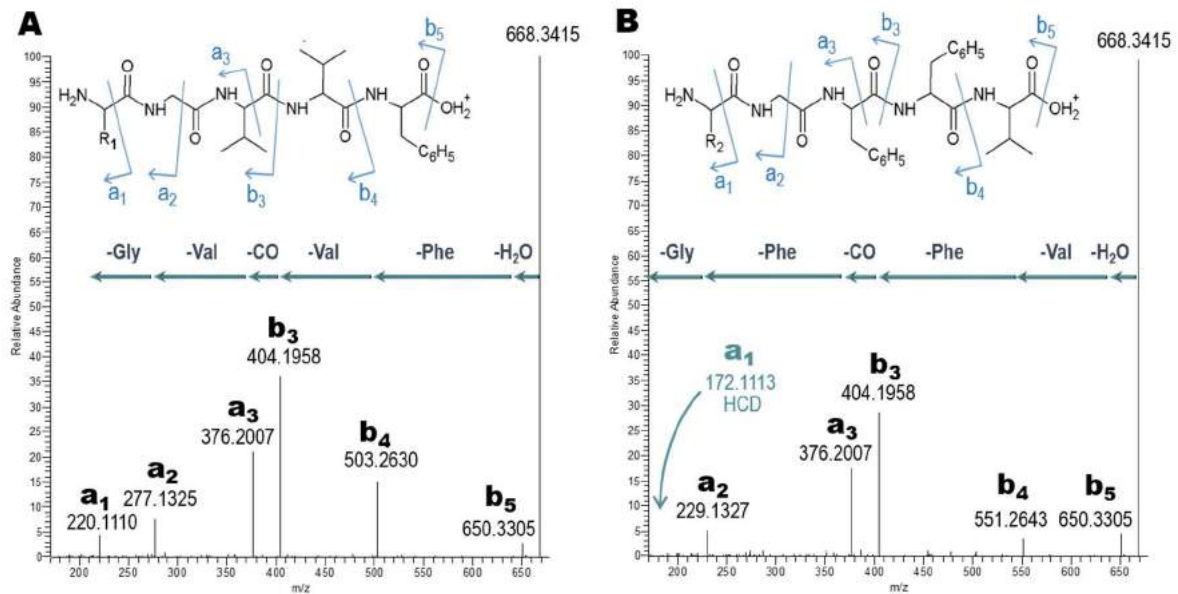
### Funding Sources

This study was co-funded by INRA and French Minister of Higher Education and Research in the context of a project supported by French National Agency of Research (ANR-15-CE21-0010-21 Newmyco).

### Supplementary data



**Fig. S1.** Annotation of the HCD spectra of fungisporin D  $C_{30}H_{37}N_5O_4$  detected at  $m/z$  532.2901 at 36.4 min (positive ionization).



**Fig. S2.** Interpretation of the high-resolution CID spectra (HCD when specified) of the new protonated peptides detected at  $m/z$  668.3415 at 24.2 min (A) and 23.8 min (B).

## Travail expérimental – La production de fungisporines par *P. nordicum*

	molecular formula <sup>a</sup>	parent ion	<sup>12</sup> C m/z	R <sub>t</sub> (min)	activation mode (%)	m/z of majors fragment ions (% relative abundance in HCD mode - % relative abundance in CID mode)
	C <sub>5</sub> H <sub>9</sub> NO	[M+H] <sup>+</sup>	100.0755	2.9	HCD(35) CID(25)	100.0753(100-100)
(9)	C <sub>6</sub> H <sub>10</sub> O <sub>4</sub>	[M+H] <sup>+</sup>	147.0649	3.1	HCD(20) CID(25)	147.0640(5-10) 129.0542(100-100) 117.0543(30-72) 97.0280(15-5) 85.0280(15-6) 99.0437(8-2) 87.0437(8-2) 75.0437(6-2) 71.0488(5-0) 55.0017(5-0) 115.0387(3-1)
	C <sub>10</sub> H <sub>13</sub> N <sub>2</sub> O <sub>2</sub>	[M+H] <sup>+</sup>	180.1016	3.9	HCD(20) CID(25)	180.1016(100-12) 121.0644(95-0) 138.0909(22-70) 162.0909(0-1)
(6)	C <sub>11</sub> H <sub>16</sub> O <sub>4</sub>	[M+H] <sup>+</sup>	213.1120	15.6	HCD(20) CID(25)	213.1119(100-45) 195.1012(1-50) 111.0438(0-100) 167.1062(0-90) 139.0386(0-27) 153.0906(0-24) 181.0854(0-20) 125.0229(0-17) 107.0853(0-12) 141.0541(0-11) 135.0800(0-10)
	C <sub>11</sub> H <sub>16</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	229.1071	5.5	HCD(20) CID(15)	229.1071(100-4) 211.0960(95-100) 141.0542(35-35) 125.0229(30-5) 157.0491(10-7) 193.0858(6-3)
	C <sub>11</sub> H <sub>16</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	229.1071	6.1	HCD(20) CID(15)	229.1071(70-45) 211.0693(100-100) 141.0544(27-15) 152.0231(25-2) 157.0494 5-2) 193.0856(4-0) 197.0806(0-1)
(7)	C <sub>11</sub> H <sub>18</sub> O <sub>4</sub>	[M+H] <sup>+</sup>	215.1277	13.9	HCD(20) CID(15)	215.1276(27-40) 197.1170(100-100) 153.1272(15-2) 137.0958(7-1) 169.1220(4-0) 151.1115(4-0) 183.1013(3-2) 119.0853(2-0) 155.1063(2-0) 165.0907(2-0)
(8)	C <sub>11</sub> H <sub>18</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	231.1226	5.7	HCD(20) CID(25)	231.1227(100-42) 181.0857(70-100) 127.0387(31-70) 213.1119(25-40) 195.1013(25-25) 81.0332(32-21) 155.1064(19-40) 139.0750(19-19) 141.0543(14-65) 151.1115(13-6) 195.1014(0-25)
	C <sub>12</sub> H <sub>9</sub> N <sub>3</sub> O <sub>2</sub>	[M-H] <sup>-</sup>	226.0620	15.2	HCD(50) CID(25)	226.0620(100-22) 196.0701(12-100) 145.0413(9-16) 183.0569(0-72) 199.0515(0-20) 198.0682(0-15)
	C <sub>12</sub> H <sub>9</sub> N <sub>3</sub> O <sub>2</sub>	[M-H] <sup>-</sup>	226.0620	12.7	HCD(50) CID(25)	226.0620(100-23) 183.0566(7-100) 196.0702(0-47) 145.0412(0-40) 199.0516(0-25)
	C <sub>12</sub> H <sub>9</sub> N <sub>3</sub> O <sub>2</sub>	[M-H] <sup>-</sup>	226.0620	18.0	HCD(50) CID(25)	226.0620(100-10) 196.0702(20-100) 183.0569(0-42) 145.0413(0-20) 199.0513(0-10) 195.0670(0-10) 198.0682(0-6)
	C <sub>13</sub> H <sub>17</sub> NO	[M-H] <sup>-</sup>	202.1236	12.8	HCD(50) CID(25)	202.1236(100-20) 131.0540(4-100) 103.0559(3-55) 158.0616(0-12) 175.1133(0-8) 124.0772(0-7)
	C <sub>14</sub> H <sub>26</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	275.1849	18.1	HCD(20) CID(25)	275.1849(15-2) 109.1009(100-3) 257.1744(55-100) 221.1532(60-4) 239.1639(42-7) 127.1114(40-5) 161.1321(22-0) 203.1425(15-0) 67.0539(15-0) 175.1476(7-0) 95.0852(4-0)
	C <sub>15</sub> H <sub>24</sub> O <sub>2</sub>	[M+H] <sup>+</sup>	237.1847	17.2	HCD(20) CID(15)	237.1847(72-7) 219.1740(100-100) 205.0854(5-5) 163.1113(0-6) 173.0591(5-3) 161.1322(0-3) 191.1167(2-2) 201.1634(0-2) 179.1062(0-1)
	C <sub>15</sub> H <sub>24</sub> O <sub>2</sub>	[M+H] <sup>+</sup>	237.1847	22.1	HCD(20) CID(15)	237.1847(100-75) 219.1739(10-100) 177.1633(3-30) 109.1009(5-25) 121.1007(0-21) 201.1633(0-21) 95.0852(6-19) 181.1216(0-19) 137.1320(0-17) 191.1788(0-12) 167.1062(0-11)
	C <sub>15</sub> H <sub>24</sub> O <sub>2</sub>	[M+H] <sup>+</sup>	237.1847	23.1	HCD(20) CID(15)	237.1847(100-75) 219.1738(12-100) 177.1633(0-32) 109.1009(5-26) 121.1008(0-24) 95.0851(5-22) 137.1320(0-19) 181.1217(0-19) 201.1634(0-19) 191.1789(0-15) 167.1062(0-13)
	C <sub>15</sub> H <sub>24</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	253.1797	9.7	HCD(20) CID(15)	253.1794(55-3) 235.1689(100-100) 217.1583(12-6) 193.1217(0-3) 177.1269(5-3) 151.0749(8-1) 189.1633(3-1)
	C <sub>15</sub> H <sub>24</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	253.1797	22.1	HCD(20) CID(15)	253.1794(100-50) 235.1689(52-100) 207.1739(35-52) 179.1062(15-19) 137.0957(7-10) 189.1633(7-15) 217.1583(5-15) 165.0906(5-5) 175.1476(5-10) 191.1790(0-2)
		[M-H] <sup>-</sup>	251.1650		HCD(50) CID(15)	251.1650(100-100) 207.1758(0-25) 137.0975(0-12) 236.0689(0-5)
	C <sub>15</sub> H <sub>24</sub> O <sub>3</sub>	[M-H] <sup>-</sup>	251.1650	35.5	HCD(50) CID(25)	251.1650(100-3) 205.1602(0-100) 233.1553(0-50) 223.1706(0-5)
	C <sub>15</sub> H <sub>24</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	253.1797	23.2	HCD(20) CID(15)	253.1795(100-2) 235.0853(50-100) 207.1735(35-62) 179.1060(29-31) 217.1584(0-14) 189.1634(0-16) 175.1477(0-11) 137.0958(0-7) 165.0906(0-4)
		[M-H] <sup>-</sup>	251.1650		HCD(50) CID(15)	251.165(100-100) 207.1758(0-30) 137.0675(0-7) 236.0695(0-4)
	C <sub>15</sub> H <sub>26</sub> O <sub>2</sub>	[M-H] <sup>-</sup>	237.1857	40.4	HCD(50) CID(25)	237.1857(100-3) 168.1148(15-100) 196.1096(0-17) 210.1252(0-14)
(5)	C <sub>16</sub> H <sub>11</sub> N <sub>3</sub> O <sub>2</sub>	[M+H] <sup>+</sup>	278.0922	9.9	HCD(20) CID(15)	278.0922(100-40) 249.0657(0-100) 235.0865(0-12) 250.0970(0-4) 146.0234(0-4) 233.0707(0-2) 175.0498(0-2) 132.0441(0-1) 105.0333(0-1)
	C <sub>16</sub> H <sub>11</sub> N <sub>3</sub> O <sub>2</sub>	[M+H] <sup>+</sup>	278.0922	15.1	HCD(35) CID(25)	278.0922(100-100) 250.0972(1-60) 233.0706(0-30) 249.0656(1-14) 248.0815(0-13) 235.0863(0-10) 132.0440(0-6) 260.0813(0-3) 222.1022(0-1)
	C <sub>16</sub> H <sub>14</sub> O <sub>3</sub>	[M-H] <sup>-</sup>	253.0862	11.5	HCD(50) CID(25)	253.0862(100-3) 159.0457(0-100) 235.0770(0-82)
	C <sub>17</sub> H <sub>10</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	295.0595	11.2	CID(25)	295.0596(60) 238.0620(100) 250.0858(25) 267.0650(20) 250.0585(17) 223.0748(15) 278.0804(16) 205.0643(13) 251.0697(11) 239.0696(10) 120.0804(6)
		[M-H] <sup>-</sup>	293.0450		HCD(50) CID(25)	293.0450(100-2) 249.0561(0-100)
	C <sub>17</sub> H <sub>14</sub> N <sub>2</sub> O <sub>3</sub>	[M-H] <sup>-</sup>	293.0923	11.3	CID(25)	293.0923(3) 249.0562(100)
	C <sub>17</sub> H <sub>26</sub> O <sub>4</sub>	[M-2H+Na] <sup>-</sup>	315.1571	38.6	CID(25)	315.1571(2) 211.0365(100) 210.0330(80) 239.1446(75) 167.0273(20) 298.0723(19) 168.0306(19) 255.1970(5) 271.1702(5)
	C <sub>17</sub> H <sub>34</sub> O <sub>5</sub>	[M-H] <sup>-</sup>	317.2320	39.1	HCD(50) CID(25)	317.2320(100-2) 257.2129(13-100) 299.2222(0-25) 255.2332(0-15) 213.0336(0-4) 212.0301(0-3) 115.0769(0-4)
	C <sub>17</sub> H <sub>34</sub> O <sub>5</sub>	[M-H] <sup>-</sup>	317.2320	39.5	HCD(50) CID(25)	317.2330(100-2) 257.2126(9-100) 299.2219(0-12) 255.233(0-7)
	C <sub>17</sub> H <sub>34</sub> O <sub>5</sub>	[M-H] <sup>-</sup>	317.2320	41.2	HCD(50) CID(25)	317.2310(100-2) 257.2129(15-100) 299.223(0-23) 255.2333(0-16) 115.0770(0-5) 195.1757(0-4)
	C <sub>17</sub> H <sub>34</sub> O <sub>5</sub>	[M-H] <sup>-</sup>	317.2320	38.1	HCD(50) CID(25)	317.2320(100-3) 257.2129(23-100) 299.2225(0-18) 273.2591(0-20) 255.2333(0-15) 115.0769(0-5) 219.2119(0-4) 195.1758(0-3)
	C <sub>17</sub> H <sub>34</sub> O <sub>5</sub>	[M-H] <sup>-</sup>	317.2320	41.0	HCD(50) CID(25)	317.2320(100-4) 257.2128(14-100) 299.2236(2-22) 255.2333(0-16) 115.0769(0-5) 195.1756(0-3)
(4)	C <sub>18</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	341.1606	15.2	HCD(20) CID(25)	341.1601(60-3) 324.1337(100-100) 296.1388(10-2) 151.1114(3-0) 195.1012(1-0)
		[M-H] <sup>-</sup>	339.1458		HCD(50) CID(25)	339.1458(100-2) 268.1097(50-100) 183.0124(0-10) 284.0918(0-4) 324.0638(0-5)
	C <sub>18</sub> H <sub>20</sub> N <sub>4</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	341.1606	20.1	HCD(20) CID(25)	341.1605(57-3) 324.1338(100-100) 296.1389(7-2)
(3)	C <sub>18</sub> H <sub>22</sub> N <sub>4</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	343.1761	12.7	HCD(20) CID(25)	343.1758(45-2) 326.1496(100-100) 298.1546(5-1)
	C <sub>18</sub> H <sub>22</sub> N <sub>4</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	343.1761	17.8	HCD(20) CID(15)	343.1756(45-49) 326.1494(100-100) 298.1544(6-1) 270.1544(3-0)
	C <sub>18</sub> H <sub>22</sub> N <sub>4</sub> O <sub>4</sub>	[M+Na] <sup>+</sup>	381.1521	15.3	HCD(35) CID(15)	381.1521(100-45) 280.0687(50-56) 298.0792(30-100) 279.0848(22-44) 364.1255(4-42) 363.1424(12-25) 169.0369(5-0) 281.0526(0-30) 254.0894(0-11) 337.1626(0-5) 261.1226(0-3)
	C <sub>18</sub> H <sub>24</sub> N <sub>2</sub> O <sub>2</sub>	[M+H] <sup>+</sup>	301.1900	37.7	HCD(35) CID(25)	301.1908(100-47) 283.1809(5-72) 227.1178(4-100) 258.1358(1-62) 240.1260(0-71) 228.1623(0-45) 255.1860(0-23) 257.2017(0-16) 199.1231(0-13) 256.1938(0-11) 212.1308(0-6)
	C <sub>19</sub> H <sub>37</sub> NO <sub>2</sub>	[M+H] <sup>+</sup>	312.2896	31.8	HCD(20) CID(15)	312.2890(6-1) 294.2788(100-100) 86.0597(6-0) 276.2685(4-3) 259.2418(4-4) 277.2528(2-4) 264.2682(2-0)

## Travail expérimental – La production de fungisporines par *P. nordicum*

<b>(1)</b>	<b>C<sub>20</sub>H<sub>18</sub>NO<sub>6</sub>Cl</b>	[M+H] <sup>+</sup>	404.0894	33.2	HCD(20) CID(15)	404.0891(31-70) 358.0836(100-100) 239.0102(65-3) 341.0571(12-2) 257.0208(11-0) 120.0804(11-0) 386.0784(4-25)
		[M-H] <sup>-</sup>	402.0750		HCD(50) CID(25)	402.0750(37-2) 358.0856(100-100) 211.0171(10-1) 254.0229(6-2) 322.1088(4-3) 270.0696(4-4) 340.0750(3-3) 239.0122(3-0) 296.0852(2-1)
<b>(2)</b>	<b>C<sub>20</sub>H<sub>19</sub>NO<sub>6</sub></b>	[M+H] <sup>+</sup>	370.1283	26.5	HCD(20) CID(25)	370.1283(32-2) 205.0491(100-17) 324.1225(37-100) 223.0597(7-0) 352.1171(7-20) 307.0959(7-5) 120.0804(7-5)
		[M-H] <sup>-</sup>	368.1138		HCD(50) CID(25)	368.1138(45-2) 324.1247(100-100) 220.0621(8-4) 280.1347(3-2) 164.0722(3-0) 306.1141(0-1)
	<b>C<sub>20</sub>H<sub>30</sub>O<sub>4</sub></b>	[M-H] <sup>-</sup>	333.2062	39.4	HCD(50) CID(25)	333.2062(100-2) 163.0768(0-100) 289.2175(0-93) 315.1962(0-40) 235.1344(0-25) 273.2070(0-22) 292.1840(0-16) 318.1993(0-14) 145.0661(0-7)
	<b>C<sub>21</sub>H<sub>30</sub>N<sub>4</sub>O<sub>4</sub></b>	[M-H] <sup>-</sup>	401.2183	35.4	HCD(50) CID(25)	401.2183(100-2) 262.1561(0-100) 172.1093(0-67) 357.1342(0-47) 245.1299(0-45) 228.1028(0-16) 203.0828(0-15) 383.2090(0-9) 137.0723(0-9) 236.1407(0-9) 214.1552(0-4)
	<b>C<sub>21</sub>H<sub>30</sub>N<sub>4</sub>O<sub>4</sub></b>	[M-H] <sup>-</sup>	401.2183	25.8	HCD(50) CID(25)	401.2183(100-3) 172.1094(12-70) 262.1562(0-100) 245.1300(0-45) 383.2092(0-21) 203.0828(0-22) 228.1033(0-22) 214.1562(0-18) 137.0723(0-12) 283.1817(0-12) 236.1406(0-12)
	<b>C<sub>21</sub>H<sub>32</sub>N<sub>2</sub>O<sub>2</sub></b>	[M+H] <sup>+</sup>	345.2519	27.0	HCD(35) CID(15)	345.2532(100-100) 245.2009(30-75) 244.1931(17-37) 274.1670(10-30) 174.1147(7-23) 330.2292(2-5) 231.1851(0-4) 285.2316(0-1) 328.1549(0-1) 327.1943(0-2)
	<b>C<sub>21</sub>H<sub>32</sub>N<sub>2</sub>O<sub>2</sub></b>	[M+H] <sup>+</sup>	345.2519	23.5	HCD(35) CID(15)	345.2532(100-100) 245.2009(20-25) 244.1932(15-13) 274.1673(8-12) 174.1149(5-9) 330.2298(3-3) 149.1071(1-0) 231.1852(1-1) 327.2423(0-1) 285.2319(0-1)
	<b>C<sub>21</sub>H<sub>32</sub>N<sub>4</sub>O<sub>5</sub></b>	[M-H] <sup>-</sup>	419.2286	17.1	CID(25)	419.2286(3) 186.0563(100) 259.1455(40) 389.1575(25)
	<b>C<sub>21</sub>H<sub>36</sub>O<sub>3</sub></b>	[M+H] <sup>+</sup>	337.2723	37.1	HCD(20) CID(25)	337.2742(100-50) 263.2372(6-100) 161.1326(0-75) 175.1481(0-72) 147.1168(0-69) 245.2267(3-67) 133.1012(0-63) 121.1011(0-59) 163.1482(0-59) 149.1325(0-55) 337.2742(0-51)
	<b>C<sub>22</sub>H<sub>28</sub>N<sub>4</sub>O<sub>5</sub></b>	[M-2H+Na] <sup>-</sup>	449.1793	25.4	CID(25)	449.1808(2) 202.0599(100) 282.1249(73)
	<b>C<sub>23</sub>H<sub>26</sub>N<sub>2</sub>O<sub>4</sub></b>	[M-2H+Na] <sup>-</sup>	415.1635	17.9	CID(25)	415.1640(2) 268.0957(100) 246.1139(65)
	<b>C<sub>25</sub>H<sub>32</sub>N<sub>4</sub>O<sub>5</sub></b>	[M+H] <sup>+</sup>	469.2409	21.2	HCD(20) CID(15)	469.2430(12-2) 332.1961(100-100) 205.0966(40-10) 148.0752(8-1) 177.1016(5-0) 265.1540(4-1) 100.1116(1-0)
	<b>C<sub>25</sub>H<sub>32</sub>N<sub>4</sub>O<sub>5</sub></b>	[M+H] <sup>+</sup>	469.2409	24.1	HCD(20) CID(15)	469.2430(23-2) 332.1961(100-100) 205.0966(55-16) 148.0752(15-2) 177.1016(6-2) 265.1540(5-2) 100.1117(1-0) 441.1897(2-0)
	<b>C<sub>25</sub>H<sub>40</sub>N<sub>4</sub>O<sub>5</sub></b>	[M+H] <sup>+</sup>	477.3060	16.9	HCD(20) CID(15)	477.3060(30-2) 265.1550(100-65) 213.1592(45-16) 185.1642(36-6) 312.2272(30-26) 459.2964(0-10) 431.3003(0-8) 166.0858(5-3) 72.0804(5-0) 86.0960(5-0) 378.2373(1-4) 120.0804(2-0) 267.2079(2-4)
	<b>C<sub>26</sub>H<sub>42</sub>O<sub>7</sub></b>	[M+H] <sup>+</sup>	467.2980	30.5	HCD(20) CID(15)	467.2999(12-2) 305.2472(100-100) 109.0646(25-0) 323.2574(16-20) 287.2370(12-12) 209.1896(8-4) 149.1323(7-0) 449.2884(0-16) 183.1013(0-5) 269.2263(0-4) 365.2682(0-18)
	<b>C<sub>28</sub>H<sub>36</sub>N<sub>4</sub>O<sub>6</sub></b>	[M+H] <sup>+</sup>	525.2682	25.8	HCD(20) CID(15)	525.2680(15-2) 346.2115(100-26) 426.2016(80-80) 247.1436(60-5) 497.2742(0-100) 327.1331(55-6) 299.1383(55-0) 378.2013(40-7) 152.0701(25-0) 279.1334(22-2) 120.0803(22-0)
		[M-H] <sup>-</sup>	523.2543		HCD(50) CID(15)	523.2543(20-100) 401.2204(100-8)
<b>(15)</b>	<b>C<sub>28</sub>H<sub>38</sub>N<sub>4</sub>O<sub>5</sub></b>	[M+H] <sup>+</sup>	511.2906	18.0	HCD(20) CID(15)	511.2906(10-100) 265.1542(100-85) 219.1488(35-10) 247.1437(25-13) 120.0804(16-0) 394.2120(8-22) 412.2228(7-4) 295.1437(3-0) 349.1905(3-6) 493.2799(0-6) 465.2876(0-8) 72.0808(2-0)
		[M-H] <sup>-</sup>	509.2770		HCD(50) CID(15)	509.2770(100-100) 276.1033(93-0) 116.0722(83-0) 115.0882(32-0) 164.0722(25-0) 202.1242(15-0) 263.1404(18-0) 262.1565(13-0) 246.1140(8-0) 146.0616(12-0) 465.2884(6-0)
<b>(21)</b>	<b>C<sub>28</sub>H<sub>38</sub>N<sub>4</sub>O<sub>5</sub></b>	[M+H] <sup>+</sup>	511.2906	18.7	HCD(20) CID(15)	511.2906(15-100) 265.1542(100-64) 219.1489(66-15) 247.1437(55-35) 346.2118(35-60) 120.0803(10-0) 72.0804(8-0) 301.1904(7-15) 166.0858(6-4) 493.2799(0-4) 394.2116(0-2)
		[M-H] <sup>-</sup>	509.2770		HCD(50) CID(15)	509.2770(73-100) 164.0722(100-0) 465.2881(15-0) 147.0457(16-0) 219.1508(7-0) 263.1406(7-0) 361.2253(4-0) 318.2193(4-0) 262.1565(5-0)
	<b>C<sub>28</sub>H<sub>38</sub>N<sub>4</sub>O<sub>6</sub></b>	[M+H] <sup>+</sup>	527.2850	12.7	HCD(20) CID(15)	527.2853(10-100) 265.1541(100-100) 235.1437(32-9) 263.1387(21-12) 120.0803(11-0) 410.2070(7-30) 428.2172(7-3) 136.0753(5-0) 365.1852(3-8) 219.1487(3-0) 311.1384(3-0) 509.2742(0-6)
		[M-H] <sup>-</sup>	525.2718		HCD(50) CID(15)	525.2718(90-100) 292.0982(100-0) 116.0722(70-0) 278.1515(20-0) 172.1095(15-0) 263.1406(13-0) 248.1088(9-0) 162.0567(7-0) 375.2406(5-0) 142.0666(7-0)
	<b>C<sub>28</sub>H<sub>42</sub>O<sub>5</sub></b>	[M+Na] <sup>+</sup>	481.2905	43.8	HCD(35) CID(15)	481.2905(100-100) 355.2243(21-55)
	<b>C<sub>28</sub>H<sub>42</sub>O<sub>5</sub></b>	[M+Na] <sup>+</sup>	481.2905	44.2	HCD(35) CID(15)	481.2905(100-100) 355.2243(12-35)
	<b>C<sub>28</sub>H<sub>42</sub>N<sub>4</sub>O<sub>5</sub></b>	[M+Na] <sup>+</sup>	537.3050	24.4	HCD(20) CID(15)	537.3050(25-2) 245.1636(100-38) 273.1587(75-88) 372.2281(40-100)
		[M-2H+Na] <sup>-</sup>	535.2920		CID(35)	535.2904(3) 491.2972(100) 448.2603(60) 362.2082(90)
	<b>C<sub>28</sub>H<sub>42</sub>N<sub>4</sub>O<sub>5</sub></b>	[M+Na] <sup>+</sup>	537.3050	23.7	HCD(20) CID(15)	537.3050(30-2) 273.1587(80-100) 372.2270(0-60) 245.1636(100-60) 420.2267(2-40) 265.1539(0-15)
		[M-2H+Na] <sup>-</sup>	535.2920		CID(35)	535.2904 (2) 276.1028(100) 232.1131 (3) 410.2078 (3)
	<b>C<sub>28</sub>H<sub>42</sub>N<sub>4</sub>O<sub>5</sub></b>	[M+Na] <sup>+</sup>	537.3050	23.1	HCD(20) CID(15)	537.3050(20-1) 245.1636(100-62) 273.1587(70-100) 265.1539(25-30) 420.2273(15-70) 438.2379(8-8) 146.0960(7-0) 519.2935(0-12) 491.3002(0-10) 120.0803(4-0) 174.1269(0-8)
		[M-2H+Na] <sup>-</sup>	535.2920		CID(15)	535.2910(2) 410.2088(3) 276.1028(100) 232.1132(10)
	<b>C<sub>29</sub>H<sub>38</sub>N<sub>4</sub>O<sub>6</sub></b>	[M+H] <sup>+</sup>	539.2860	26.9	HCD(20) CID(15)	539.2860(15-2) 120.0803(100-0) 265.1537(35-20) 478.0348(13-0) 422.2057(0-100) 521.2787(0-12) 511.2900(0-10) 374.2055(0-5) 351.2137(0-4) 412.2214(0-4)
		[M-H] <sup>-</sup>	537.2698		CID(15)	537.2698(100)
<b>(13)</b>	<b>C<sub>29</sub>H<sub>40</sub>N<sub>4</sub>O<sub>5</sub></b>	[M+H] <sup>+</sup>	525.3046	19.5	HCD(20) CID(15)	525.3051(13-12) 265.1541(100-100) 261.1594(32-20) 120.0803(15-0) 412.2226(10-2) 219.1487(11-2) 408.2278(8-22) 233.1658(30-10) 507.2984(0-5) 479.3033(0-3) 363.2077(1-10) 295.1453(3-1) 86.0966(2-0)
<b>(12)</b>	<b>C<sub>29</sub>H<sub>40</sub>N<sub>4</sub>O<sub>5</sub></b>	[M+H] <sup>+</sup>	525.3046	20.2	HCD(20) CID(15)	525.3060(7-2) 279.1698(100-100) 219.1487(27-12) 247.1437(24-30) 120.0804(15-0) 394.2118(10-10) 426.2382(7-3) 479.3037(0-4) 349.1926(2-10) 295.1455(1-1) 233.1661(1-5) 72.0804(1-0)
<b>(10)</b>	<b>C<sub>29</sub>H<sub>40</sub>N<sub>4</sub>O<sub>6</sub></b>	[M+H] <sup>+</sup>	541.3039	15.4	HCD(20) CID(15)	541.3039(7-20) 279.1697(100-100) 235.1436(30-5) 263.1384(20-12) 410.2068(11-30) 120.0803(9-0) 442.2329(6-4) 136.0752(6-0) 365.1851(3-6) 523.2893(0-7) 495.2946(0-4) 311.1380(3-0) 233.1656(1-4)
		[M-H] <sup>-</sup>	539.2860		HCD(50) CID(15)	539.2860(100-100) 292.0982(95-1) 130.0877(90-0)
<b>(14)</b>	<b>C<sub>29</sub>H<sub>40</sub>N<sub>4</sub>O<sub>6</sub></b>	[M+H] <sup>+</sup>	541.3039	13.8	HCD(20) CID(15)	541.3039(10-2) 265.1541(100-100) 249.1592(37-12) 277.1540(30-23) 120.0803(12-0) 428.2170(10-0) 424.2226(10-32) 523.2889(0-10) 379.1998(0-8) 311.1383(4-0) 219.1490(2-3) 495.2946(0-4) 283.1453(2-0) 136.0760(3-0) 86.0966(3-0)
		[M-H] <sup>-</sup>	539.2860		HCD(50) CID(15)	539.2860(70-100) 292.0980(100-2) 116.0724(55-0)
	<b>C<sub>29</sub>H<sub>50</sub>N<sub>2</sub>O<sub>3</sub></b>	[M+Na] <sup>+</sup>	497.3741	43.6	HCD(20) CID(15)	497.3741(100-90) 217.1335(25-100) 245.2012(1-5) 174.1152(0-12) 159.0718(0-12) 479.3641(0-9) 439.3324(0-5)
	<b>C<sub>29</sub>H<sub>51</sub>N<sub>2</sub>O<sub>2</sub></b>	[M+H] <sup>+</sup>	446.3966	38.6	HCD(35) CID(25)	446.3994(50-2) 429.3721(100-100) 177.0905(38-16) 163.0749(19-0) 137.0593(9-0) 428.3897(0-10) 428.2668(0-9) 387.3259(0-5) 400.2719(0-4) 411.3629(0-3) 417.3731(0-2)
<b>(20)</b>	<b>C<sub>30</sub>H<sub>37</sub>N<sub>5</sub>O<sub>4</sub></b>	[M+H] <sup>+</sup>	532.2901	36.4	HCD(20) CID(15)	532.2908(75-2) 433.2228(100-82) 504.2947(35-100) 385.2229(90-19) 334.1541(84-28) 159.0913(75-0) 258.1595(62-0) 286.1544(60-4) 346.2117(53-7) 306.1594(35-0) 120.0804(15-0)

## Travail expérimental – La production de fungisporines par *P. nordicum*

		[M-H] <sup>-</sup>	530.2750		CID(15)	530.2755(100) 401.2197(4)
(22)	C <sub>30</sub> H <sub>37</sub> N <sub>5</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	548.2845	29.4	HCD(20) CID(15)	548.2866(45-3) 385.2227(100-28) 449.2172(70-85) 350.1491(62-25) 286.1544(61-4) 258.1595(55-2) 159.0912(55-0) 362.2067(48-10) 520.2906(27-100) 322.1546(25-0) 503.2634(0-20)
		[M-H] <sup>-</sup>	546.2707		HCD(50) CID(15)	546.2707(35-100) 417.2151(100-0) 311.1732(35-0)
	C <sub>30</sub> H <sub>37</sub> N <sub>5</sub> O <sub>6</sub>	[M-H] <sup>-</sup>	562.2649	26.2	HCD(50) CID(15)	562.2662(45-100) 311.1729(100-0) 440.2313(45-0) 297.1569(20-0) 433.2106(20-0) 323.1730(12-0)
(19)	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	550.3015	17.4	HCD(20) CID(15)	550.3015(10-95) 304.1649(100-100) 287.1384(35-62) 159.0913(24-1) 433.2230(7-30) 219.1487(7-1) 451.2333(5-2) 247.1436(5-1) 334.1541(4-2) 388.2039(2-10) 532.2935(0-15) 504.2982(0-2) 259.1454(0-1) 120.0811(2-0)
		[M-H] <sup>-</sup>	548.2875		HCD(50) CID(15)	548.2875(100-100) 116.0722(90-0) 315.1142(50-0) 186.0564(45-0) 164.0722(37-0) 419.2308(22-0) 259.1457(25-0) 202.0991(15-0) 375.2407(10-0) 302.1514(20-0) 351.0327(5-0)
(17)	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	550.3015	19.1	HCD(20) CID(15)	550.3015(25-100) 286.1544(100-80) 258.1594(96-32) 265.1540(65-30) 385.2228(50-90) 241.1330(12-14) 364.2222(7-0) 269.1278(6-1) 199.1436(6-0) 166.0858(6-0) 340.2012(1-12) 72.0803(6-0) 533.2770(1-5) 532.2930(0-2) 213.1392(2-0)
		[M-H] <sup>-</sup>	548.2875		HCD(50) CID(15)	548.2875(52-100) 164.0721(100-0) 202.0990(42-0) 375.2406(17-0) 147.0456(15-0) 301.1675(7-0) 504.2989(5-0) 400.2363(4-0) 263.1404(4-0)
(18)	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	[M+H] <sup>+</sup>	566.2959	12.3	HCD(20) CID(15)	566.2965(7-62) 304.1649(100-100) 287.1384(32-68) 159.0913(23-0) 235.1436(10-0) 449.2173(7-35) 467.2280(6-6) 263.1384(7-2) 404.1964(4-10) 548.2852(0-8) 350.1491(4-1) 520.2904(0-2) 136.0751(2-0) 190.1217(1-0) 259.1454(0-2)
		[M-H] <sup>-</sup>	564.2824		HCD(50) CID(15)	564.2824(100-100) 116.0722(95-0) 202.0514(90-0) 331.1091(45-0) 435.2256(30-0) 119.0507(33-0) 302.1514(20-0) 285.1936(18-0) 278.1513(17-0) 318.1464(15-0)
(16)	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	[M+H] <sup>+</sup>	566.2959	13.2	HCD(20) CID(15)	566.2965(25-25) 286.1544(100-88) 385.2228(57-100) 258.1595(99-35) 281.1489(67-38) 241.1330(15-15) 340.2009(0-12) 380.2170(5-1) 520.2939(0-2) 549.2725(0-3) 548.2885(0-2) 213.1388(2-0) 159.0922(3-0) 130.0655(3-0) 72.0808(3-0)
		[M-H] <sup>-</sup>	564.2824		HCD(50) CID(15)	564.2824(50-100) 180.0669(100-0) 202.0993(30-0)
	C <sub>30</sub> H <sub>43</sub> N <sub>5</sub> O <sub>5</sub>	[M+Na] <sup>+</sup>	576.3165	21.0	HCD(20) CID(15)	576.3165(75-35) 312.1699(100-100) 411.2381(72-90) 284.1750(47-25) 185.1067(42-12) 265.1539(35-12) 158.0960(31-0) 254.1496(10-0) 364.2224(7-0) 196.0751(7-0) 558.3054(0-6) 530.3108(0-5) 72.0804(7-0)
		[M-2H+Na] <sup>-</sup>	574.3030		HCD(50) CID(15)	574.3033(40-100) 164.0721(100-0)
	C <sub>30</sub> H <sub>43</sub> N <sub>5</sub> O <sub>6</sub>	[M+Na] <sup>+</sup>	592.3107	15.5	HCD(20) CID(15)	592.3107(15-32) 261.1585(100-30) 304.1646(90-100) 289.1538(80-60) 287.1381(43-70) 475.2326(13-85) 159.0912(9-0) 493.2435(6-5) 574.3008(0-20) 411.2384(5-3) 162.0907(5-0) 312.1703(5-2) 546.3048(0-4) 241.1328(0-1)
		[M-H] <sup>-</sup>	578.2984		CID(15)	578.2984(100)
	C <sub>31</sub> H <sub>41</sub> N <sub>5</sub> O <sub>5</sub>	[M-H] <sup>-</sup>	562.3030	19.2	CID(15)	562.3036(100)
	C <sub>31</sub> H <sub>41</sub> N <sub>5</sub> O <sub>5</sub>	[M-H] <sup>-</sup>	562.3030	21.4	CID(15)	562.3035(100)
	C <sub>31</sub> H <sub>42</sub> O <sub>6</sub>	[M-H] <sup>-</sup>	531.2722	37.3	CID(15)	531.2715(100)
	C <sub>31</sub> H <sub>45</sub> N <sub>5</sub>	[M+Na] <sup>+</sup>	510.3554	36.9	HCD(20) CID(10)	510.3561(10-100) 377.3194(100-15)
	C <sub>31</sub> H <sub>45</sub> N <sub>5</sub>	[M+Na] <sup>+</sup>	510.3554	38.2	HCD(20) CID(10)	510.3561(20-100) 377.3193(100-5) 393.3388(0-20)
	C <sub>32</sub> H <sub>40</sub> O <sub>6</sub>	[M+H] <sup>+</sup>	521.2866	39.5	HCD(20) CID(15)	521.2875(100-35) 131.0851(33-0) 475.2835(32-100) 205.1218(30-11) 299.1634(19-20) 471.2483(0-35) 485.2636(0-20) 461.2634(0-20) 443.2575(11-34) 489.2627(11-60) 503.2785(10-75)
	C <sub>32</sub> H <sub>42</sub> O <sub>7</sub>	[M-H] <sup>-</sup>	537.2838	39.5	HCD(50) CID(15)	537.2845(10-100) 71.0507(100-0) 95.0507(60-0) 147.0821(40-0) 417.2808(52-0) 159.1184(27-0) 97.0663(35-0) 321.2232(15-0) 399.2707(15-0) 431.2970(20-0) 177.1290(12-0)
	C <sub>32</sub> H <sub>42</sub> O <sub>8</sub>	[M-H] <sup>-</sup>	553.2810	38.3	HCD(50) CID(15)	553.2810(10-100) 121.0664(100-0) 97.0300(95-0) 175.1133(80-0) 201.0926(30-0) 219.1033(17-0) 393.2444(23-0)
	C <sub>32</sub> H <sub>44</sub> O <sub>9</sub>	[M-H] <sup>-</sup>	571.2899	36.8	HCD(50) CID(15)	571.2899(4-100) 87.0456(100) 175.1133(40-0) 363.2335(20-0) 439.2496(17-0) 113.0614(23-0) 121.0663(18-0)
	C <sub>34</sub> H <sub>64</sub> O <sub>8</sub>	[M-H] <sup>-</sup>	599.4507	39.2	HCD(50) CID(15)	599.4507(2-100) 317.2339(100-0)
	C <sub>38</sub> H <sub>39</sub> N <sub>5</sub> O <sub>7</sub>	[M-H] <sup>-</sup>	676.2778	36.9	HCD(50) CID(15)	676.2778(12-100) 330.0888(100-0)
	C <sub>38</sub> H <sub>45</sub> N <sub>5</sub> O <sub>6</sub>	[M+H] <sup>+</sup>	668.3415	23.8	HCD(20) CID(15)	668.3415(17-100) 404.1962(100-25) 376.2011(62-18) 229.1330(50-12) 172.1113(4-0) 551.2631(0-2) 650.3317(0-5)
	C <sub>38</sub> H <sub>45</sub> N <sub>5</sub> O <sub>6</sub>	[M+H] <sup>+</sup>	668.3415	24.2	HCD(20) CID(15)	668.3415(23-100) 650.3305(0-4) 503.2640(21-17) 404.1963(100-38) 376.2010(75-23) 277.1328(48-9) 220.1108(0-5)

**Table S1.** Secondary metabolites detected from *P. nordicum* after culture on wheat grains. The HCD and CID spectra are presented as a list of fragment ions with their relative abundance. <sup>a</sup> The chemical formulas were calculated after comparison of the *m/z* ratios obtained from the culture of *P. nordicum* in <sup>12</sup>C wheat, 97% <sup>13</sup>C wheat and 53% <sup>13</sup>C / 96% <sup>15</sup>N wheat.

## References

- [1] S. Brase, A. Encinas, J. Keck, C.F. Nising, Chemistry and biology of mycotoxins and related fungal metabolites, *Chem. Rev.* 109 (2009) 3903–3990.
- [2] M.C. Fisher, D.A. Henk, C.J. Briggs, J.S. Brownstein, L.C. Madoff, S.L. McCraw, S.J. Gurr, Emerging fungal threats to animal, plant and ecosystem health, *Nature*. 484 (2013) 1–18.
- [3] E.M. Fox, B.J. Howlett, Secondary metabolism: Regulation and role in fungal biology, *Curr. Opin. Microbiol.* 11 (2008) 481–487.
- [4] Food and Agriculture Organization of the United Nations, Mycotoxin regulations in 2003 and current developments, 2003. (2016).
- [5] T. Kuiper-Goodman, P.M. Scott, Risk assessment of the mycotoxin ochratoxin A, *Biomed. Environ. Sci.* 2 (1989) 179–248.
- [6] International Agency for Research on Cancer (IARC), Ochratoxin A, *IARC Monogr. Eval. Carcinog. Risk Chem. to Humans*. 56 (1993) 489–521.
- [7] Y. Wang, L. Wang, F. Liu, Q. Wang, J.N. Selvaraj, F. Xing, Y. Zhao, Y. Liu, Ochratoxin A producing fungi, biosynthetic pathway and regulatory mechanisms, *Toxins (Basel)*. 8 (2016) 1–15.
- [8] P. Kroghl, B. Haidt, R. Pleftinas, S. Ceovic, Balkan (endemic) nephropathy and foodborn ochratoxin A: Preliminary results of a survey of foodstuffs, *Acta Path. Microbiol. Scand. Sect. B*. 85 (1977) 238–240.
- [9] F. Malir, V. Ostry, A. Pfohl-Leszkwicz, J. Malir, J. Toman, Ochratoxin A: 50 years of research, *Toxins (Basel)*. 8 (2016) 12–15.
- [10] T. Bui-Klimke, F. Wu, Evaluating weight of evidence in the mystery of balkan endemic nephropathy, *Risk Anal.* 34 (2014) 1688–1705.
- [11] F. Wu, T. Bui-Klimke, K. Naumoff Shields, Potential economic and health impacts of ochratoxin A regulatory standards, *World Mycotoxin J.* 7 (2014) 387–398.
- [12] F.J. Cabañes, M.R. Bragulat, G. Castellá, Ochratoxin A producing species in the genus *Penicillium*, *Toxins (Basel)*. 2 (2010) 1111–1120.
- [13] J.C. Frisvad, R.A. Samson, Polyphasic taxonomy of *Penicillium* subgenus *Penicillium*: A guide to identification of food and air-borne terverticillate *Penicillia* and their mycotoxins, *Stud. Mycol.* 49 (2004) 1–174.
- [14] M.T. Henke, N.L. Kelleher, Modern mass spectrometry for synthetic biology and structure-based discovery of natural products, *Nat. Prod. Rep.* 33 (2016) 942–950.
- [15] W.B. Dunn, A. Erban, R.J.M. Weber, D.J. Creek, M. Brown, R. Breitling, T. Hankemeier, R. Goodacre, S. Neumann, J. Kopka, M.R. Viant, Mass appeal: Metabolite identification in mass spectrometry-focused untargeted metabolomics, *Metabolomics*. 9 (2013) 44–66.
- [16] C. Bueschl, R. Krška, B. Kluger, R. Schuhmacher, Isotopic labeling-assisted metabolomics using LC-MS, *Anal. Bioanal. Chem.* 405 (2013) 27–33.
- [17] P.M. Cano, E.L. Jamin, S. Tadrict, P. Bourdaud’Hui, M. Péan, L. Debrauwer, I.P. Oswald, M. Delaforge, O. Puel, New untargeted metabolic profiling combining mass spectrometry and isotopic labeling: Application on *Aspergillus fumigatus* grown on wheat, *Anal. Chem.* 85 (2013) 8412–8420.
- [18] T. Hautbergue, O. Puel, S. Tadrict, L. Meneguetti, M. Péan, M. Delaforge, L. Debrauwer, I.P. Oswald, E.L. Jamin, Evidencing 98 secondary metabolites of *Penicillium verrucosum* using substrate isotopic labeling and high resolution mass spectrometry, *J. Chrom. B.* DOI:10.101 (2017).
- [19] R.O. Studer, Synthesis and structure of Fungisporin, *Experientia*. 25 (1969) 899.
- [20] H. Ali, M.I. Ries, P.P. Lankhorst, R.A.M. Van Der Hoeven, O.L. Schouten, M. Noga, T. Hankemeier, N.N.M.E. Van Peij, R.A.L. Bovenberg, R.J. Vreeken, A.J.M. Driessen, A non-canonical NRPS is involved in the synthesis of fungisporin and related hydrophobic cyclic tetrapeptides in *Penicillium chrysogenum*, *PLoS One*. 9 (2014) e98212.
- [21] A. Klitgaard, J.B. Nielsen, R.J.N. Frandsen, M.R. Andersen, K.F. Nielsen, Combining Stable Isotope Labeling and Molecular Networking for Biosynthetic Pathway Characterization, *Anal. Chem.* 87 (2015) 6520–6526.
- [22] J.C. Nielsen, S. Grijsseels, S. Prigent, B. Ji, J. Dainat, K.F. Nielsen, J.C. Frisvad, M. Workman, J. Nielsen, Global analysis of biosynthetic gene clusters reveals vast potential of secondary metabolite production in *Penicillium* species, *Nat. Microbiol.* 2 (2017) 17044.
- [23] Y.C.S. Adjovi, S. Bailly, B.J.G. Gnonlonfin, S. Tadrict, A. Querin, A. Sanni, I.P. Oswald, O. Puel, J.D. Bailly, Analysis of the contrast between natural occurrence of toxigenic *Aspergilli* of the *Flavi* section and aflatoxin B1 in cassava, *Food Microbiol.* 38 (2014) 151–159.
- [24] T.O. Larsen, A. Svendsen, J. Smedsgaard, Biochemical Characterization of Ochratoxin A-Producing Strains of the Genus *Penicillium*, *Appl. Environ. Microbiol.* 67 (2001) 3630–3635.
- [25] M. Pean, S. Boiry, J.-C. Ferrandi, F. Gibiat, O. Puel, M. Delaforge, Production and use of mycotoxins uniformly enriched with stable isotopes for their dosage in biological samples: (1) Production of uniformly enriched biomass, *J. Label. Compd. Radiopharm.* 50 (2007) 563–564.
- [26] L.W. Sumner, A. Amberg, D. Barrett, M.H. Beale, R. Beger, C.A. Daykin, T.W. Fan, O. Fiehn, R. Goodacre, J.L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI), *NIH*. 3 (2007) 211–221.
- [27] F. Allen, A. Pon, M. Wilson, R. Greiner, D. Wishart, CFM-ID: A web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra, *Nucleic Acids Res.* 42 (2014) 1–6.
- [28] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape:

- a software environment for integrated models of biomolecular interaction networks, *Genome Res.* 13 (2003) 2498–2504.
- [29] G.M. Leite, N. Magan, Á. Medina, Comparison of different bead-beating RNA extraction strategies: An optimized method for filamentous fungi, *J. Microbiol. Methods.* 88 (2012) 413–418.
- [30] M. Röttig, M.H. Medema, K. Blin, T. Weber, C. Rausch, O. Kohlbacher, NRPSpredictor2 - A web server for predicting NRPS adenylation domain specificity, *Nucleic Acids Res.* 39 (2011) 362–367.
- [31] K.F. Nielsen, T.O. Larsen, N. Keller, T.S. Bugni, The importance of mass spectrometric dereplication in fungal secondary metabolite analysis, *Front. Microbiol.* 6 (2015) 71.
- [32] T. Kind, O. Fiehn, Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry, *BMC Bioinformatics.* 8 (2007) 105.
- [33] H. Laatsch, *Antibase 2012: The natural compound identifier*, (2012). <https://application.wiley-vch.de/stmdata/antibase.php> (accessed January 2, 2017).
- [34] T.O. Larsen, H. Franzzyk, S.R. Jensen, UV-guided isolation of verrucines A and B, novel quinazolines from *Penicillium verrucosum* structurally related to anacine from *Penicillium aurantiogriseum*, *J. Nat. Prod.* 62 (1999) 1578–1580.
- [35] K.F. Nielsen, J. Smedsgaard, Fungal metabolite screening: Database of 474 mycotoxins and fungal metabolites for dereplication by standardised liquid chromatography-UV-mass spectrometry methodology, *J. Chromatogr. A.* 1002 (2003) 111–136.
- [36] L. Rahbaek, S. Sperry, J.C. Frisvad, T.O. Larsen, PC-2, LL-P888gamma and some novel analogue alpha-pyrones from *Penicillium nordicum*, *P. verrucosum* and *P. olsonii*, *Biochem. Syst. Ecol.* 31 (2003) 313–317.
- [37] W. Demuth, M. Karlovits, K. Varmuza, Spectral similarity versus structural similarity: Mass spectrometry, *Anal. Chim. Acta.* 516 (2004) 75–85.
- [38] M. Wang, J.J. Carver, V. V Phelan, L.M. Sanchez, N. Garg, Y. Peng, D.D. Nguyen, J. Watrous, C.A. Kapon, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A. V Melnik, M.J. Meehan, W.-T. Liu, M. Crüsemann, P.D. Boudreau, E. Esquenazi, M. Sandoval-Calderón, R.D. Kersten, L.A. Pace, R.A. Quinn, K.R. Duncan, C.-C. Hsu, D.J. Floros, R.G. Gavilan, K. Kleigrewe, T. Northen, R.J. Dutton, D. Parrot, E.E. Carlson, B. Aigle, C.F. Michelsen, L. Jelsbak, C. Sohlenkamp, P. Pevzner, A. Edlund, J. McLean, J. Piel, B.T. Murphy, L. Gerwick, C.-C. Liaw, Y.-L. Yang, H.-U. Humpf, M. Maansson, R.A. Keyzers, A.C. Sims, A.R. Johnson, A.M. Sidebottom, B.E. Sedio, A. Klitgaard, C.B. Larson, C.A. Boya P, D. Torres-Mendoza, D.J. Gonzalez, D.B. Silva, L.M. Marques, D.P. Demarque, E. Pociute, E.C. O’Neill, E. Briand, E.J.N. Helfrich, E.A. Granatosky, E. Glukhov, F. Ryffel, H. Houson, H. Mohimani, J.J. Kharbush, Y. Zeng, J.A. Vorholt, K.L. Kurita, P. Charusanti, K.L. McPhail, K.F. Nielsen, L. Vuong, M. Elfeki, M.F. Traxler, N. Engene, N. Koyama, O.B. Vining, R. Baric, R.R. Silva, S.J. Mascuch, S. Tomasi, S. Jenkins, V. Macherla, T. Hoffman, V. Agarwal, P.G. Williams, J. Dai, R. Neupane, J. Gurr, A.M.C. Rodríguez, A. Lamsa, C. Zhang, K. Dorrestein, B.M. Duggan, J. Almaliti, P.-M. Allard, P. Phapale, L.-F. Nothias, T. Alexandrov, M. Litaudon, J.-L. Wolfender, J.E. Kyle, T.O. Metz, T. Peryea, D.-T. Nguyen, D. VanLeer, P. Shinn, A. Jadhav, R. Müller, K.M. Waters, W. Shi, X. Liu, L. Zhang, R. Knight, P.R. Jensen, B.Ø. Palsson, K. Pogliano, R.G. Lington, M. Gutiérrez, N.P. Lopes, W.H. Gerwick, B.S. Moore, P.C. Dorrestein, N. Bandeira, Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking, *Nat. Biotechnol.* 34 (2016) 828–837.
- [39] A.M. Frank, M.E. Monroe, A.R. Shah, J.J. Carver, N. Bandeira, R.J. Moore, G.A. Anderson, R.D. Smith, P.A. Pevzner, Spectral archives: extending spectral libraries to analyze both identified and unidentified spectra, *Nat. Methods.* 8 (2011) 587–91.
- [40] S.M. Williams, J.S. Brodbelt, MSn characterization of protonated cyclic peptides and metal complexes, *J Am Soc Mass Spectrom.* 15 (2004) 1039–1054.
- [41] K.F. Medzihradzsky, R.J. Chalkley, Lessons in de novo peptide sequencing by tandem mass spectrometry, *Mass Spectrom. Rev.* 34 (2015) 43–63.
- [42] C. Jia, W. Qi, Z. He, B. Qiao, Sequencing peptides by electrospray ion-trap mass spectrometry: A useful tool in synthesis of Axinastatin 3, *Cent. Eur. J. Chem.* 4 (2006) 620–631.
- [43] B. Paizs, S. Suhal, Fragmentation pathways of protonated peptides, *Mass Spectrom. Rev.* 24 (2005) 508–548.
- [44] R.J. Chalkley, K.C. Hansen, M.A. Baldwin, Bioinformatic methods to exploit mass spectrometric data for proteomic applications, *Methods Enzymol.* 402 (2005) 289–312.
- [45] K. Biemann, Nomenclature for peptide fragment ions (positive ions), *Methods Enzymol.* 193 (1990) 886–887.
- [46] J.B. Parker, C.T. Walsh, Stereochemical outcome at four stereogenic centers during conversion of prephenate to tetrahydrotyrosine by BacABGF in the bacilysin pathway, *Biochemistry.* 51 (2012) 5622–5632.
- [47] B. V. Bertinetti, N.I. Peña, G.M. Cabrera, An antifungal tetrapeptide from the culture of *Penicillium canescens*, *Chem. Biodivers.* 6 (2009) 1178–1184.
- [48] J.I. Pitt, *Penicillium viridicatum*, *Penicillium verrucosum*, and production of ochratoxin A, *Appl. Environ. Microbiol.* 53 (1987) 266–269.



## **2. ARTICLE 4. UN MÉCANISME DE FRAGMENTATION PEPTIDIQUE NON RÉFÉRENCÉ**

---

Dans de l'étude des métabolomes secondaires de *P. verrucosum* et *P. nordicum*, les mécanismes classiques de fragmentation peptidiques ont été appliqués pour l'élucidation structurale de peptides linéaires. Des expériences de fragmentations successives par MS<sup>3</sup> et MS<sup>4</sup> ont été réalisées sur l'un des peptides linéaires connus (WVVF).

Lors de l'interprétation de l'arbre de fragmentation du peptide WVVF, des ions non référencés dans la littérature concernant la fragmentation peptidique ont été observés. Des mécanismes ont donc été proposés afin d'expliquer ces potentiels nouveaux mécanismes. Ces observations ont été vérifiées et précisées par l'analyse MS<sup>n</sup> d'autres tétrapeptides dont les standards étaient disponibles au laboratoire. La comparaison de leurs spectres MS<sup>n</sup> en modes CID et HCD a permis d'étudier la filiation des ions fragments.

Cette section présente l'ébauche d'un article qui pourra faire l'objet d'une communication dans un journal de spectrométrie de masse ou d'un article à part entière lorsque les résultats seront enrichis par des analyses complémentaires. En effet, afin de vérifier les nouveaux mécanismes secondaires proposés, des analyses par échange hydrogène/deutérium ainsi que des calculs théoriques quantiques permettant d'évaluer les états de transitions sont nécessaires. Une partie de la discussion (discussion, section 1.4) détaille ces perspectives. Ces analyses traitent de spectrométrie de masse d'un aspect fondamental. C'est pourquoi ces résultats sont présentés dans un article distinct de l'Article 3 présentant l'analyse du métabolome secondaire de *P. nordicum*.





## Evidencing a new peptidic fragmentation mechanism

Thaïs Hautbergue<sup>a,b</sup>, Olivier Puel<sup>a</sup>, Jean-Claude Tabet<sup>c,d</sup>, Laurent Debrauwer<sup>a,b</sup>, Isabelle P. Oswald<sup>a</sup>, Emilien L. Jamin<sup>a,b,\*</sup>

<sup>a</sup> Toxalim (Research Centre in Food Toxicology), Université de Toulouse, INRA, ENVT, INP-Purpan, UPS, Toulouse, France

<sup>b</sup> Axiom platform, MetaToul-MetaboHUB, National Infrastructure for Metabolomics and Fluxomics, Toulouse, France

<sup>c</sup> CEA, iBiTec-S, SPI, LEMM, Metabohub Gif Sur Yvette, France

<sup>d</sup> Université Paris VI (UPMC), CNRS UMR 7201, Paris Cedex 05, France

\* Corresponding author: emilien.jamin@toulouse.inra.fr

**ABSTRACT:** Structural elucidation of tetrapeptides by successive mass spectrometry experiments in high resolution (HPLC-ESI-HRMS) has highlighted the presence of some product ions never referenced. First, the presence of these product ions were verified in the product ion spectra of 11 tetrapeptides. Then, a fragmentation mechanism was propose based on MS<sup>3</sup> and MS<sup>4</sup> experiments.

### 1. Introduction

Tandem mass spectrometry experiments by high-resolution mass spectrometry have shown that some product ions non referenced were generated during the fragmentation of tetrapeptides. This article proposes a fragmentation mechanism bases on MS<sup>n</sup> experiments.

### 2. Material and Methods

#### 2.1. Chemicals and Reagents.

The formic acid as well as the solvents (acetonitrile, ethyl acetate and methanol) came from Fisher Scientific (Thermo Fisher Scientific, Illkirch, France). The solvent used for sample preparation were HPLC grade. Acetonitrile was MS grade. Ultra-pure water was generated from a Milli-Q system (Millipore, Saint Quentin en Yvelines, France). The peptides were ordered from GeneCust (GeneCust, Dudelange, Luxembourg). All standards were dissolved in methanol for qualitative analysis by HPLC-MS and the solutions were kept at -20°C.

#### 2.2. Liquid Chromatography and Mass Spectrometry Analyses

The peptides were analyzed coupling HPLC with HRMS. The chromatographic separation was performed with an Ultimate 3000 system (Thermo Fisher Scientific, Les Ulis, France). The phase A of the mobile phase contained 0.1% formic acid in water and the phase B contained 100% acetonitrile. The two solvents were eluted at the flow rate of 0.2 mL/min according to the following gradient: 0 min 20% of B, 30 min 50% of B, from 35 to 45 min 90% of B, from 50 to 60 20% of B. A volume of 10 µL of the injected sample was eluted with a Luna C18 column (150 mm × 2.0 mm, 5 µm) (Phenomenex, Torrance, CA, U.S.A.). HRMS analyses were performed with a LTQ Orbitrap XL hybrid high-resolution mass spectrometer (Thermo Fisher Scientific, Les Ulis, France) in positive and negative electrospray ionization (ESI) modes. Ionization parameters were as follows: spray voltage +4.5 kV, capillary temperature 350 °C, sheath gas (N<sub>2</sub>) flow rate 40 a.u. (arbitrary units), auxiliary gas (N<sub>2</sub>) flow rate 6 a.u. for the positive mode, and spray voltage -3.7 kV, capillary temperature 350 °C, sheath gas flow rate

30 a.u. (N<sub>2</sub>), auxiliary gas flow rate 10 a.u. (N<sub>2</sub>) for the negative mode. High-resolution mass spectra were acquired at a resolution of 60,000 from  $m/z$  50 to  $m/z$  800. UV Analysis of the fungal extract was also performed after a chromatographic separation on an Ultimate 3000 system (Thermo Fisher Scientific, Les Ulis, France). The high-resolution mass analyzer was calibrated in each ionization mode based on the supplier's protocol and calibration mixtures (Thermo Fisher Scientific, Les Ulis, France).

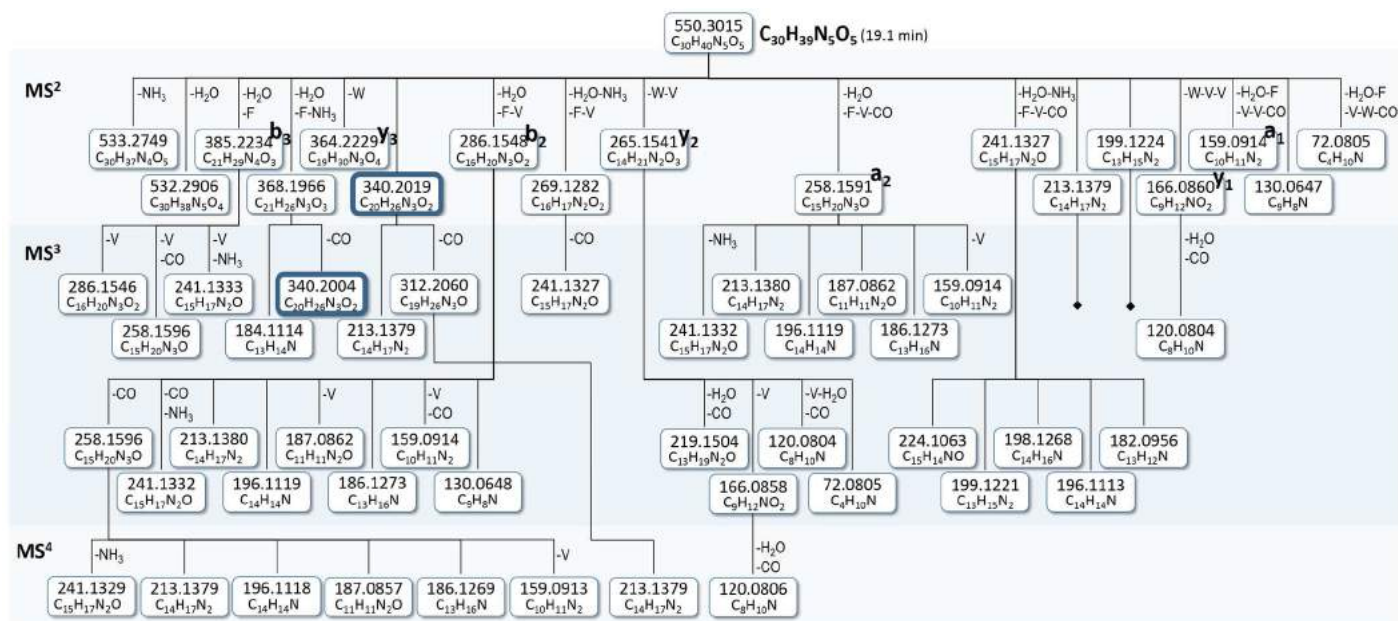
Targeted high-resolution MS/MS experiments were performed in both resonant and non-resonant modes on each peptide. In this article, CID refers to MS fragmentations in resonant processes and HCD refers to non-resonant fragmentations according to the nomenclature used by ThermoQuest. The HCD mode was used due to the absence of low mass cut-off and pseudopotential well. HCD and CID modes provide complementary structural information. Whereas CID mode lead to competitive fragmentations, HCD mode lead to consecutive fragmentation. [Parameters for MS/MS by resonant CID mode: resolution = 7500, MS<sup>n</sup> maximum ion time = 200 ms, isolation width = 1.5  $m/z$ , Act.  $q_z$  = 0.250 and activation time = 30 ms]. [Parameters for MS/MS by higher energy collision dissociation (HCD): resolution = 7500, last MS<sup>n</sup> step maximum ion time = 200 ms, isolation width = 1.5  $m/z$  and activation time = 30 ms]. HCD and CID normalized collision energies (NCE) were adjusted to obtain appropriate fragmentation (10% of survival parent ion abundance) and maximum structural information.

### 3. Results

#### 3.1. MS<sup>n</sup> Experiments for the Analysis of WVVF

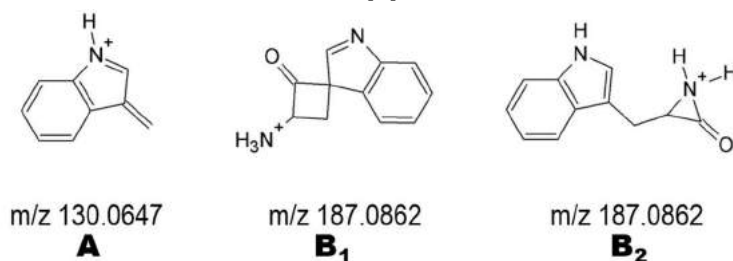
Peptidic fragmentation mechanisms were investigated by MS<sup>n</sup> experiments on the linear tetrapeptide WVVF. This peptide has been chosen for two reasons. Firstly, it corresponds to a secondary metabolite of the fungisporin family produced by *P. verrucosum* and studied in our laboratory [1]. Secondly, it allows taking into account the potential influence of the N terminal tryptophan, which is known to be involved in specific fragmentation mechanisms such as losses of ammoniac [2,3]. Successive fragmentations were performed in HCD mode in order to allow the detection of low  $m/z$  ions (Figure 1).

Interpretation of this fragmentation pattern is performed following classical mechanisms of peptidic fragmentations [5][6]. Thus, ions of **a**, **b** and **y** series are highlighted as well as their satellite ions resulting from losses of amino acid residues (147.0698 for a phenylalanine, 99.0688 for a valine and 186.0786 for a tryptophan), and/or by consecutive losses of CO, H<sub>2</sub>O. These classical fragmentations are annotated by their respective losses in Figure 1. Additional product ions which arise from the loss of NH<sub>3</sub> are also detected, according to the specific fragmentations generated by N terminal tryptophan [2,3]. This loss are explained by energy resolved CID experiments demonstrating that this fragmentation has the lowest energy threshold in case of protonated tryptophan [3].



**Fig. 1.** Fragmentation tree of the tetrapeptide WVVF ( $C_{30}H_{39}N_5O_5$ ). For each  $MS^n$  experiment,  $MS^{n-1}$  fragmentations were performed in CID mode in the linear trap (NCE 25%), and the  $n^{\text{th}}$  fragmentation was performed by HCD activation mode (NCE 20%). Black points illustrate the absence of fragmentation of an ion at this energy. Edges illustrate neutral losses between an ion and its fragments and are annotated according to classical fragmentation mechanisms [4]. The two ions framed in bold are the subject of a discussion in this article.

Alternatively, several product ions are not annotated by classical neutral losses. Among them, the product ion detected at  $m/z$  130.0647 corresponds to the indole ring doubly linked to the carbon beta of the tryptophan, as already referenced [3] (Figure 2A). The ion observed at  $m/z$  187.0862 is expected to be also specific of peptides with a N terminal tryptophan. Two structures are proposed for this product ion (Figure 2B<sub>1</sub> and 2B<sub>2</sub>). Firstly, this product ion could result from the attack of the indole ring on the carboxyl group (Figure 2B<sub>1</sub>). This mechanism have been already described for histidine [6]. Secondly, this ion could also display an aziridinone structure resulting from a nucleophile attack of the N terminal amine on the carboxyl (Figure 2B<sub>2</sub>), but it is expected that such aziridinone would be instable and would lose CO to stabilize as immonium [6].

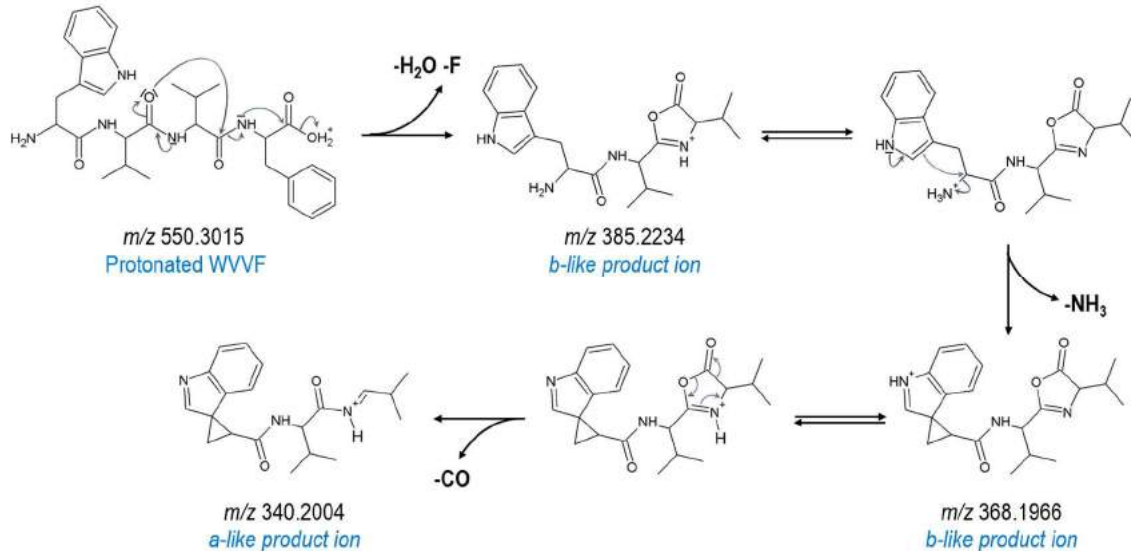


**Fig. 2.** Structures proposed for the product ions detected at  $m/z$  130.0647 (A) and  $m/z$  187.0862 (B) in the  $MS^2$ ,  $MS^3$  and  $MS^4$  spectra from the protonated WVVF.

Other product ions in Figure 1 were neither referenced nor explained by classical fragmentation mechanisms. These ions are of two types: those likely resulting from a product ion having lost ammonia from the N terminal tryptophan (ions linked with the ion  $m/z$  241.1327 for example, such as ions at  $m/z$  196.1119 and  $m/z$  199.1221), and those likely independent from the nature of the N terminal amino acid. This article focus on this later type of product ions since they could concern all peptides regardless their sequence and therefore could be useful in terms of global comprehension of peptidic fragmentation mechanisms.

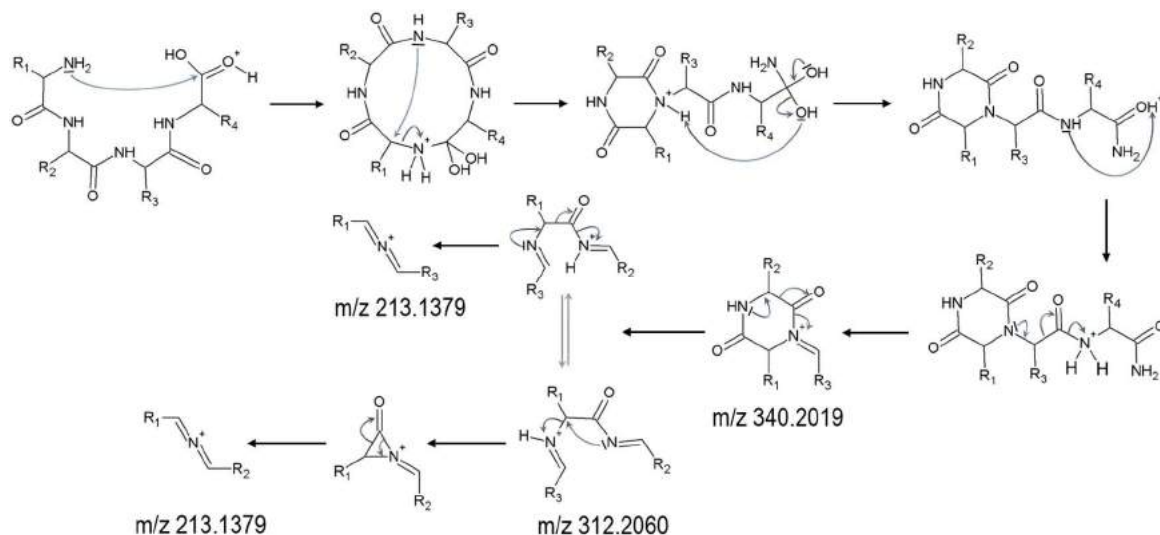
3.2. Analysis of the product ions at  $m/z$  340.2229

At first sight, the product ion  $C_{21}H_{26}N_3O_3$  detected twice in the fragmentation tree of WVVF at  $m/z$  340.2004 in  $MS^3$  and  $m/z$  340.2019 in  $MS^2$  experiments (Figure 1, ions framed in bold) are annotated as resulting from consecutive losses of water, phenylalanine, ammonia and carbon monoxide according to classical fragmentation mechanisms. This hypothesis is particularly suited to the annotation of the product ion at  $m/z$  340.2004 in the  $MS^3$  spectra from 368.1966 since this later resulted from these classical neutral losses (Figure 3).



**Fig. 3.** Fragmentation mechanism leading to the product ion at  $m/z$  340.2004 from the ion at  $m/z$  368.1966. Loss of phenylalanine residue is illustrated by “-F”.

However, this hypothesis seems not adapted for the ion at  $m/z$  340.2019 detected in the  $MS^2$  of WVVF. Indeed, the  $MS^3$  spectra of 340.2019 displays two product ions at  $m/z$  213.1379 and  $m/z$  312.2060, which could not result from the structure of the ion 340.2004 presented in Figure 3. Therefore, another fragmentation mechanism is proposed to annotate these three ions (Figure 4).



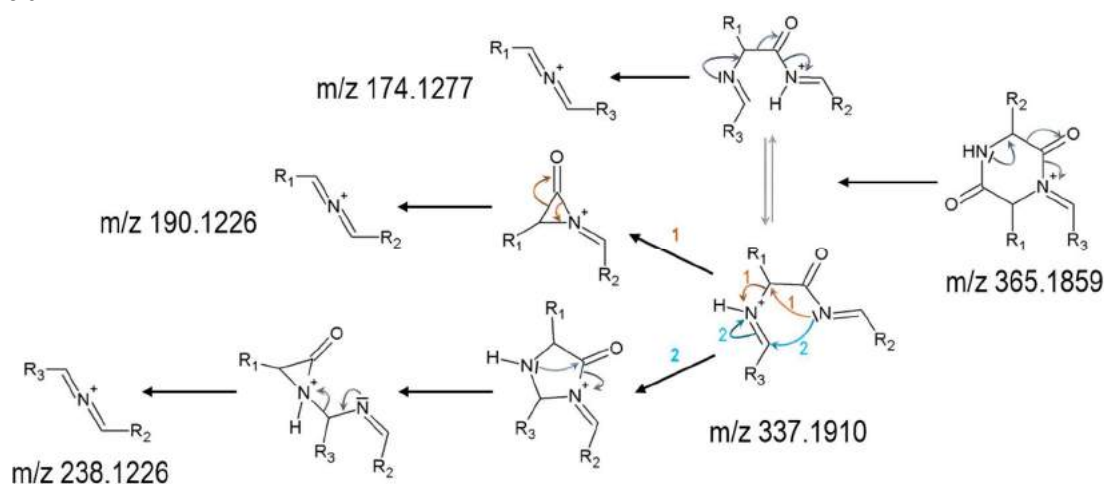
**Fig. 4.** Fragmentation mechanism proposed for the generation of the ion detected at  $m/z$  340.2019 in the  $MS^2$  spectra of the protonated WVVF and for the generation of the product ions at  $m/z$  312.2060 and  $m/z$  213.1379 in the  $MS^3$  and  $MS^4$  spectra of 340.2019.  $R_1$  represents the side chain of tryptophan,  $R_2$  and  $R_3$  represent the side chains of the valines in second and third position respectively and  $R_4$  represents the side chain of phenylalanine.

The protonation of the precursor ion on the oxygen of the carboxylic function (Figure 4) may lead to a cyclisation process by attack of the free doublet of the amine function on the electrophile carbon. Since this phenomenon is not observed for other carbonyl groups, the protonation of the C terminal carbonyl group is suggested to be promoted by the electro-attracting hydroxyl group in alpha increasing the electropositivity of the carbon. Following the cyclisation step, a reorganization may occur for the creation of a ring with six apices and containing the tryptophan and a valine residue ( $R_1$  and  $R_2$  in Figure 4). The loss of the fourth amino acid (phenylalanine,  $R_4$  in Figure 4) is then generated. To the best of our knowledge, this is the first time that such mechanism was reported.

### 3.3. A Fragmentation Mechanism Leading to Five Product Ions

The fragmentation mechanism proposed in Figure 4 do not involve the indol ring of the tryptophan, suggesting that it could also occur with other tetrapeptides that do not contain tryptophan. To support this hypothesis,  $MS^3$  and  $MS^4$  experiments have been performed on the tetrapeptide VYFV and the product ions equivalent to these observed at  $m/z$  340.2019,  $m/z$  312.2060 and  $m/z$  213.1379 in the  $MS^n$  spectra of WVVF were sought. As a result, a product ion is detected at  $m/z$  365.1859 in the HCD spectrum of the protonated tetrapeptide. This ion is likely an equivalent of the ion at  $m/z$  340.2019 of WVVF, with  $R_1$ ,  $R_2$  and  $R_3$  in Figure 4 corresponding this time to the side chains of valine, tyrosine and phenylalanine respectively. Similarly, an ion at  $m/z$  337.1910 is equivalent to the ion detected at  $m/z$  312.2070 in the  $MS^3$  spectrum of the ion at  $m/z$  365.1859. Finally, the equivalent of the ion at  $m/z$  213.1386 is detected at  $m/z$  174.1277 in the  $MS^4$  spectra from the previous ion at  $m/z$  337.1910 (Table 1, line 2).

However, two other product ions at  $m/z$  190.1226 and  $m/z$  238.1226 are detected in the  $MS^4$  spectrum obtained from the ion at  $m/z$  337.1910 (Table 1, line 2). These observations highlight that the fragmentation mechanism proposed in Figure 4 is more complex and a clarification regarding the fragmentation mechanism is proposed in Figure 5. Here we propose that three types of ions organized around quaternary ammoniums could be generated, each resulting from the loss of a different amino acid.



**Fig. 5.** Fragmentation mechanism proposed for the generation of five product ions from the protonated VYFV.  $R_1$ ,  $R_2$  and  $R_3$  represent the side chains of valine, tyrosine and phenylalanine respectively.

In order to support this hypothesis, this mechanism was investigated on the third tetrapeptide VYVW. By replacing  $R_3$  by the side chain of tryptophan, five ions at  $m/z$  404,  $m/z$  376,  $m/z$  213,  $m/z$  190

and  $m/z$  277 are expected. All these expected product ions were successfully observed in MS<sup>3</sup> and MS<sup>4</sup> experiments on the protonated VYVW (Table 1, line 2).

### 3.4. Application of the New Fragmentation Mechanism to Other Tetrapeptides

The previous analysis of WVVF do not allow the observation of the five distinct ions due to the redundant information resulting from the presence of two consecutive valines in second and third positions in the sequence. Indeed, since R<sub>2</sub> and R<sub>3</sub> were the same, no distinction could be made between the ions present in the upper part of the Figure 5. Moreover, the product ion equivalent to these detected at  $m/z$  238.1226 and  $m/z$  277.1335 in the MS<sup>n</sup> spectra of VYFV and VYVW respectively is not observed for WVVF, probably due to a too weak abundance and a lack of sensitivity of the analytical technique.

In order to support all the hypotheses stated in this article, product ions likely resulting from a such fragmentation mechanism have been sought in the MS<sup>2</sup>, MS<sup>3</sup> and MS<sup>4</sup> spectra of eight more tetrapeptides, and the results are presented in Table 1. As for WVVF, some theoretical product ions are not detected for some tetrapeptides (grey cells in Table 1, lines 1, 4, 5, 6, 9 and 10). Similarly to WVVF, all the tetrapeptides with same second and third amino acids, such as IFFV and VFFI, displayed redundant product ions (Table 1, lines 1, 7, 8, 9 and 10). Concerning these peptides, these redundant ions are reported twice in the Table 1 in order to highlight the equivalence between the product ions generated from the different tetrapeptides.

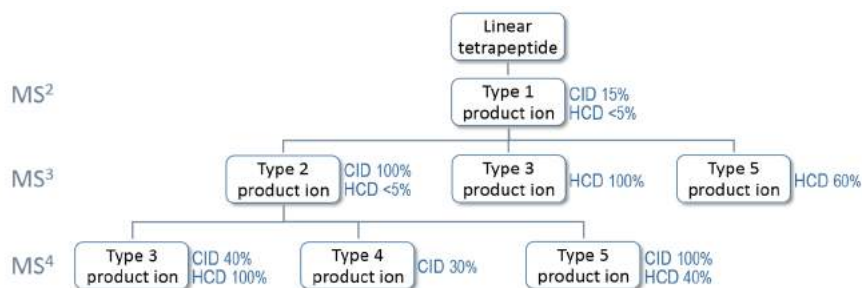
Line number	Peptide sequence	$m/z$ of the 1 <sup>st</sup> product ion	$m/z$ of the 2 <sup>nd</sup> product ion	$m/z$ of the 3 <sup>rd</sup> product ion	$m/z$ of the 4 <sup>th</sup> product ion	$m/z$ of the 5 <sup>th</sup> product ion
1	WVVF	340.2019 C <sub>20</sub> H <sub>26</sub> N <sub>3</sub> O <sub>2</sub>	312.2070 C <sub>19</sub> H <sub>26</sub> N <sub>3</sub> O	213.1386 C <sub>14</sub> H <sub>17</sub> N <sub>2</sub>	213.1386 C <sub>14</sub> H <sub>17</sub> N <sub>2</sub>	126.1277 C <sub>8</sub> H <sub>16</sub> N
2	VYFV	365.1859 C <sub>22</sub> H <sub>25</sub> N <sub>2</sub> O <sub>3</sub>	337.1910 C <sub>21</sub> H <sub>25</sub> N <sub>2</sub> O <sub>2</sub>	174.1277 C <sub>12</sub> H <sub>16</sub> N	190.1226 C <sub>12</sub> H <sub>16</sub> NO	238.1226 C <sub>16</sub> H <sub>16</sub> NO
3	VYVW	404.1968 C <sub>24</sub> H <sub>26</sub> N <sub>3</sub> O <sub>3</sub>	376.2019 C <sub>23</sub> H <sub>26</sub> N <sub>3</sub> O <sub>2</sub>	213.1386 C <sub>14</sub> H <sub>17</sub> N <sub>2</sub>	190.1226 C <sub>12</sub> H <sub>16</sub> NO	277.1335 C <sub>18</sub> H <sub>17</sub> N <sub>2</sub> O
4	VFWV	388.2019 C <sub>24</sub> H <sub>26</sub> N <sub>3</sub> O <sub>2</sub>	360.2070 C <sub>23</sub> H <sub>26</sub> N <sub>3</sub> O	213.1386 C <sub>14</sub> H <sub>17</sub> N <sub>2</sub>	174.1277 C <sub>12</sub> H <sub>16</sub> N	261.1386 C <sub>18</sub> H <sub>17</sub> N <sub>2</sub>
5	VYFI	365.1859 C <sub>22</sub> H <sub>25</sub> N <sub>2</sub> O <sub>3</sub>	337.1910 C <sub>21</sub> H <sub>25</sub> N <sub>2</sub> O <sub>2</sub>	174.1277 C <sub>12</sub> H <sub>16</sub> N	190.1226 C <sub>12</sub> H <sub>16</sub> NO	238.1226 C <sub>16</sub> H <sub>16</sub> NO
6	IYFV	379.2016 C <sub>23</sub> H <sub>27</sub> N <sub>2</sub> O <sub>3</sub>	351.2067 C <sub>22</sub> H <sub>27</sub> N <sub>2</sub> O <sub>2</sub>	188.1433 C <sub>13</sub> H <sub>18</sub> N	204.1382 C <sub>13</sub> H <sub>18</sub> NO	238.1226 C <sub>16</sub> H <sub>16</sub> NO
7	IFFV	363.2067 C <sub>23</sub> H <sub>27</sub> N <sub>2</sub> O <sub>2</sub>	335.2117 C <sub>22</sub> H <sub>27</sub> N <sub>2</sub> O	188.1433 C <sub>13</sub> H <sub>18</sub> N	188.1433 C <sub>13</sub> H <sub>18</sub> N	222.1277 C <sub>16</sub> H <sub>16</sub> N
8	VFFI	349.1899 C <sub>22</sub> H <sub>25</sub> N <sub>2</sub> O <sub>2</sub>	321.1961 C <sub>21</sub> H <sub>25</sub> N <sub>2</sub> O	174.1277 C <sub>12</sub> H <sub>16</sub> N	174.1277 C <sub>12</sub> H <sub>16</sub> N	222.1277 C <sub>16</sub> H <sub>16</sub> N
9	VFFV	349.1899 C <sub>22</sub> H <sub>25</sub> N <sub>2</sub> O <sub>2</sub>	321.1961 C <sub>21</sub> H <sub>25</sub> N <sub>2</sub> O	174.1277 C <sub>12</sub> H <sub>16</sub> N	174.1277 C <sub>12</sub> H <sub>16</sub> N	222.1277 C <sub>16</sub> H <sub>16</sub> N
10	WVYV	340.2019 C <sub>20</sub> H <sub>26</sub> N <sub>3</sub> O <sub>2</sub>	312.2070 C <sub>19</sub> H <sub>26</sub> N <sub>3</sub> O	213.1386 C <sub>14</sub> H <sub>17</sub> N <sub>2</sub>	213.1386 C <sub>14</sub> H <sub>17</sub> N <sub>2</sub>	246.1277 C <sub>18</sub> H <sub>16</sub> N

**Table 1.** Chemical composition and  $m/z$  ratios of the product ions detected in MS<sup>2</sup>, MS<sup>3</sup> and MS<sup>4</sup> spectra of 10 tetrapeptides and supporting the new fragmentation mechanism presented in Figure 4 and Figure 5. Grey cells correspond to theoretical product ions non-observed in the experimental MS<sup>n</sup> spectra.

### 3.5. HCD and CID Analyses to Study Relationship Between the Product Ions.

The fragmentations of the tetrapeptides presented in Table 1 were generated in both HCD and CID modes in order to differentiate competitive fragmentations (obtained in CID mode) from consecutive fragmentations (obtained in HCD mode). Some product ions were mostly detected in HCD mode. As they were detected at relative abundances less than 1% in the CID mode, it was difficult to perform their analysis by mass spectrometry. The increase of the activation time (from 30 ms to 60 ms)

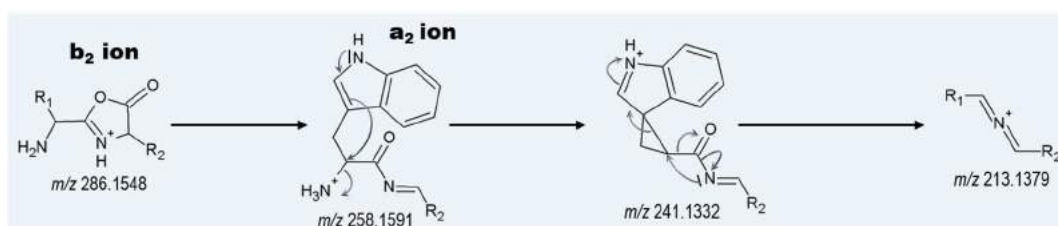
and the analysis in low resolution in the more sensitive linear trap did not improve their signal. However, the relationships between the product ions were determined by the comparison of MS<sup>n</sup> spectra generated in CID mode and in HCD mode (Figure 6).



**Fig. 6.** Relationship between the different product ions analyzed by CID and HCD mode mass spectrometry experiments. Each type of ion is labeled by a number referring to the columns of Table 1. Relative abundances are indicated in percentage.

### 3.6. Product Ions May Also Come From Other Mechanisms.

In the fragmentation tree of the protonated WVVF (Figure 1), the product ion observed at  $m/z$  213.1380 is detected five times. As it was above mentioned for the MS<sup>3</sup> spectrum from the ion at  $m/z$  340.2019, it may be generated according to the postulated fragmentation mechanism (Figure 4 and Figure 5). However, the same ions in the MS<sup>3</sup> spectrum from 286.1548 and 258.1591 and in the MS<sup>4</sup> spectrum from 258.1591 are generated by a distinct fragmentation mechanism depending on the N terminal tryptophan (Figure 7). The loss of ammonia from the ion **a**<sub>2</sub> at  $m/z$  258.1591 is allowed by the indole ring of the N terminal tryptophan, leading to a spirocyclopropane, which stabilized the structure of the ion at  $m/z$  241.1332 [2]. The product ion at  $m/z$  213.1379 resulting from the further loss of carbon monoxide consists of a tryptophan (side chain R<sub>1</sub> in Figure 7) and a valine (side chain R<sub>2</sub> in Figure 7). This structure is identical to the product ions 213.1379 presented in Figure 4 but is generated by another pathway depending on the N terminal tryptophan.



**Fig. 7.** Fragmentation mechanisms for the generation of the ions detected at  $m/z$  213.1379 in the MS<sup>3</sup> spectrum from the ions **b**<sub>2</sub> at  $m/z$  286.1548. R<sub>1</sub> = side chain of tryptophan, R<sub>2</sub> = side chains of valine.

## 4. Conclusion

*De novo* structural elucidation of the peptides by MS<sup>n</sup> experiments in both CID mode and HCD mode have led to the discovery of product ions non-referenced in the literature. A fragmentation mechanism was proposed to explain their generation and to justify their structure. Further studies are necessary to verify this hypothesis. These type of product ions could be integrated in the automatic sequencing program to help for the structural elucidation of unknown tetrapeptides. Some research would be interesting to look for corresponding product ions from peptides with more than four amino acids.



## **Acknowledgements**

PhD doctoral fellowship was co-funded by INRA and French Minister of Higher Education and Research. Authors thank the Groupe de Recherches Appliquées en Phytotechnologie, CEA, IBEB, Cadarache, Saint-Paul-les-Durance, France for labeled wheat production

The authors are grateful to Pascal Bourdaudhui who developed the software MassCompare, and to Jean-Claude Tabet for his kind help for the mass spectrometry analyses and data monitoring. All MS experiments were performed on the instruments of the MetaToul-AXIOM platform, partner of the national infrastructure of metabolomics and fluxomics: MetaboHUB (MetaboHUB-ANR-11-INBS-0010).

## **Funding Sources**

This study was co-funded by INRA and French Minister of Higher Education and Research in the context of a project supported by French National Agency of Research (ANR-15-CE21-0010-21 Newmyco).

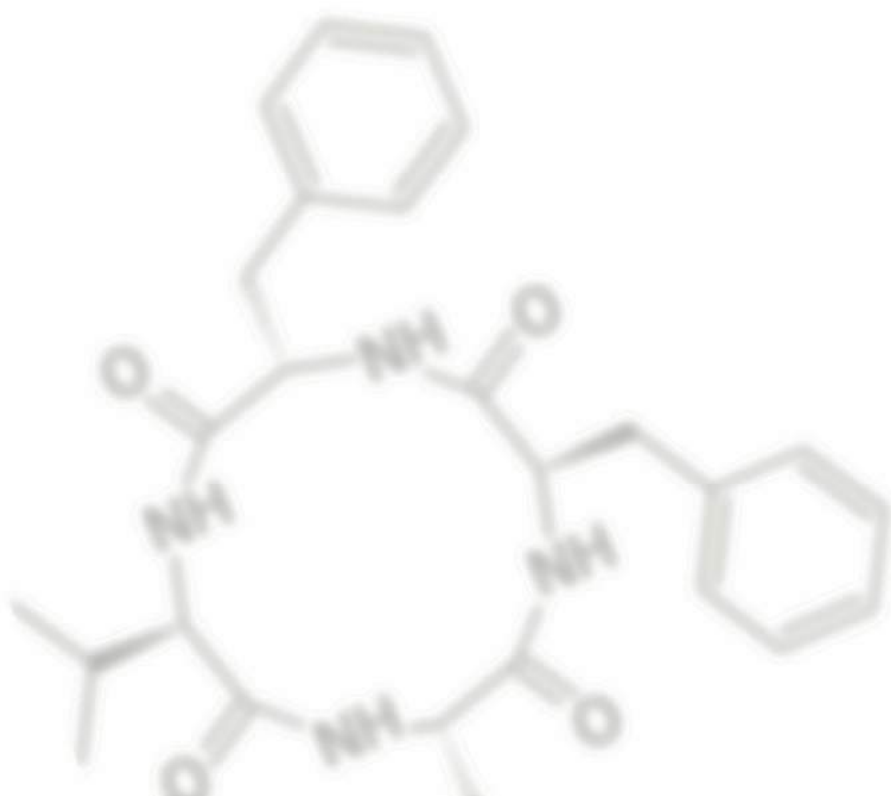
## **References**

- [1] T. Hautbergue, O. Puel, S. Tadriss, L. Meneguetti, M. Péan, M. Delaforge, L. Debrauwer, I.P. Oswald, E.L. Jamin, Evidencing 98 secondary metabolites of *Penicillium verrucosum* using substrate isotopic labeling and high resolution mass spectrometry, *J. Chrom. B.* DOI:10.101 (2017).
- [2] H. Lioe, R. a J. O’Hair, G.E. Reid, Gas phase ion Chem. biomolecules. Part 37 - Gas-phase reactions of protonated tryptophan, *J. Am. Soc. Mass Spectrom.* 15 (2004) 65–76.
- [3] H. El Aribi, G. Orlova, A.C. Hopkinson, K.W.M. Siu, Gas-phase fragmentation reactions of protonated aromatic amino acids: Concomitant and consecutive neutral eliminations and radical cation formations, *J. Phys. Chem. A.* 108 (2004) 3844–3853.
- [4] K.F. Medzihradzky, R.J. Chalkley, Lessons in de novo peptide sequencing by tandem mass spectrometry, *Mass Spectrom. Rev.* 34 (2015) 43–63.
- [5] K. Biemann, Nomenclature for peptide fragment ions (positive ions), *Methods Enzymol.* 193 (1990) 886–887.
- [6] B. Paizs, S. Suhal, Fragmentation pathways of protonated peptides, *Mass Spectrom. Rev.* 24 (2005) 508–548.

# 3

*Travail expérimental*

## **Étude du métabolome secondaire de *F. graminearum***





## 1. ARTICLE 5. LE MÉTABOLOME D'UN PUISSANT PHYTOPATHOGENE

---

Afin de comprendre les mécanismes d'infestation des champs par *F. graminearum* et de connaître les composés toxiques qu'il produit, la méthode d'étude de métabolome secondaire par double marquage isotopique du substrat fongique a également été appliquée à l'étude de *F. graminearum* au cours d'une étude précédente à cette thèse. La croissance du champignon sur grains de blé marqués a permis de détecter 37 métabolites secondaires, dont au moins 29 sont inconnus, et de connaître leur composition chimique. Des réseaux moléculaires ont été construits dans le cadre de cette thèse afin d'orienter les analyses structurales des composés inconnus et ont ainsi pu mettre en évidence la production de nouvelles mycotoxines potentielles. Contrairement à *P. verrucosum*, *F. graminearum* n'est pas saprophyte et infeste les céréales lors de leur croissance dans les champs. Les métabolites détectés sur grains de blé marqués ont donc été recherchés de manière ciblée sur des épis de blés naturellement infestés. Ce travail a également été réalisé préalablement à cette thèse. L'analyse de ces épis a montré que le blé étudié était co-infesté par *F. graminearum* et *F. culmorum*. Ce dernier est une espèce phylogéniquement proche de *F. graminearum*. Afin d'apprécier la cohérence entre proximité phylogénique et similarité métabolique, les productions de métabolites par *F. graminearum* et *F. culmorum* ont été comparées lorsqu'ils se développent dans les champs. Afin de limiter les co-infestations, des infestations artificielles ont été réalisées en cultures sous serre, au moment de la floraison. Trois souches ont été analysées : la souche de *F. graminearum* étudiée *in vitro* (souche PH1, référencée sous le code INRA 605), une souche de *F. graminearum* isolée des champs naturellement infestés (souche INRA DGEC) et une souche de *F. culmorum* isolée des champs naturellement infestés. L'Article 5, en préparation, présente les détails analytiques et les résultats obtenus lors de l'analyse du métabolome secondaire de *F. graminearum* *in vitro* et *in planta*.



## Identification of new secondary metabolites produced *in vitro* and *in planta* by *F. graminearum* on wheat using isotopic labeling and HPLC-HRMS

Patricia M. Cano <sup>a</sup>, Emilien L. Jamin <sup>a</sup>, Souria Tadriss <sup>a</sup>, Thais Hautbergue <sup>a</sup>, Pascal Bourdaudhui <sup>a</sup>, Michel Péan <sup>b,c,d</sup>, Laurent Debrauwer <sup>a</sup>, Isabelle P. Oswald <sup>a</sup>, Marcel Delaforge <sup>e</sup>, Olivier Puel <sup>a,1</sup>

<sup>a</sup> Toxalim (Research Centre in Food Toxicology), Toulouse University, INRA, ENVT, INP-Purpan, UPS, Toulouse, France

<sup>b</sup> CEA, IBEB, Groupe de Recherches Appliquées en Phytotechnologie, Cadarache, FR 13108 Saint-Paul-les-Durance, France

<sup>c</sup> CNRS, UMR Biologie Végétale & Microbiologie Environnementale, FR 13108 Saint-Paul-les-Durance, France

<sup>d</sup> Université d'Aix-Marseille, Jardin du Pharo, FR 13007 Marseille, France

<sup>e</sup> CNRS, URA 2096 and CEA Saclay, FR 91191 Gif sur Yvette, France

<sup>1</sup> Corresponding author: Olivier Puel - INRA, Toxalim, 180, Chemin de Tournefeuille, BP3, 31027 Toulouse, France - [olivier.puel@inra.fr](mailto:olivier.puel@inra.fr)

*Fusarium graminearum* is widely acknowledged as a devastating cereal pathogen and as an important producer of mycotoxins, such as deoxynivalenol, an immune disruptor, and zearalenone, an estrogenic disruptor. Therefore, contamination by *F. graminearum* makes contaminated grains unsafe for human and animal consumption. However, so far the secondary metabolome of this species has been poorly analyzed, leaving a large array of possible mycotoxins undiscovered. Here, we applied a recently validated untargeted metabolic profiling method to draw a better picture of the secondary metabolome of *F. graminearum*. This method combines high resolution mass spectrometry with isotopic labeling allowing specific and highly accurate detection of secondary metabolites as well as unambiguous formula determination. In this study, 29 new secondary metabolites could be characterized on wheat cultures *in vitro*. Among them, molecular networking highlighted two secondary metabolites similar to fusaristatin A. Based on the structure of fusaristatin A, two new members of the fusaristatin family could also be identified. Further analyses of wheat spikelets infected *in planta* with *F. graminearum* and *F. culmorum*, another important cereal pathogen, revealed, the production of many of these new metabolites in the fields. The results obtained in this study bring new perspectives on the risks related to wheat infection by these two economically important and challenging mycotoxigenic phytopathogens, encouraging towards the exploration of their metabolomes.

**Keywords** : *F. graminearum*, secondary metabolites, wheat contamination, molecular networking, fusaristatin

### INTRODUCTION

Filamentous fungi constitute one of the most important sources of secondary metabolites. So far, over 2000 fungal secondary metabolites have been described. While many of them present industrially interesting properties, others, known as mycotoxins, induce deleterious effects in humans, animals and plants[1]. Among filamentous fungi, the genus *Fusarium* is one of the most important mycotoxin-producing fungal genera[2]. *Fusarium* species are known for the production of trichothecenes, fumonisins and zearalenone which, together with aflatoxins, ochratoxin A and patulin, are the only legally regulated mycotoxins worldwide.

Among *Fusarium* species, *Fusarium graminearum* (*F. graminearum*) is one of the most devastating cereal pathogens in temperate regions of the world[3,4]. Unlike other plant pathogens, *F. graminearum* has the ability to switch from a saprophytic state in which it survives among decayed organic matter, to an endophytic state in which it develops on its host asymptotically and finally, to a pathogenic state in which it is responsible for several phytopathologies[5–7]. Contamination by *F. graminearum* is responsible for important economical losses worldwide due to yield decrease caused by head blight or scab disease in wheat and barley, and cob or ear rot disease in maize[3,7]. In addition, contamination by this fungus is also of great concern due to the production of several mycotoxins such as deoxynivalenol, zearalenone and fusarine C[8,9]. Deoxynivalenol is known for its immunosuppressive effects, zearalenone for its important estrogenic properties and fusarine C is strongly suspected of being mutagenic[10–13]. Therefore, *F. graminearum* makes contaminated cereal grains unsafe for human and animal consumption.

While many studies have addressed the pathogenicity of *F. graminearum* and the toxicity of the few known mycotoxins it produces[3,7,14–18], the secondary metabolome of this fungus remains largely unexplored. The recent sequencing of the genome of *F. graminearum* has revealed that most of the genes involved in the secondary metabolism of this fungus do not correlate with any known secondary metabolite[19]. Only 6 out of the 15 genes encoding for polyketide synthases (PKS) and only 3 out of the 19 genes encoding for non-ribosomal peptide synthases (NRPS) have been characterized in terms of their biological function[20–22]. Therefore, it is readily hypothesized that the majority of the secondary metabolites produced by *F. graminearum* remain unidentified. Gaining more knowledge on the metabolome of this fungus would not only positively affect our understanding of the phytopathologies it causes but it would also allow a better characterization of the risks related to contaminated goods. Consequently, the aim of this study was to draw a better picture of the secondary metabolites produced by *F. graminearum* on wheat, which is one of the most sensitive cereals to this fungus.

In order to fulfill this objective, a recently developed and validated untargeted detection method was used[23]. This method combines high-resolution mass spectrometry for specific detection and highly accurate mass measurement and *in vitro* stable isotopic labeling of the metabolome with carbon 13 (<sup>13</sup>C) and nitrogen 15 (<sup>15</sup>N) for formula ambiguities to be overcome by reporting the number of carbon and nitrogen atoms. One of the main advantages of this analytical method is to discriminate products with a non-biological origin, which is usually a struggling point when characterizing new metabolites. In addition, AntiBase, which is one of the most exhaustive databases for metabolites produced by microorganisms[24], was used for possible identification of the metabolites. Based on this analysis, and in order to better characterize the risks related to wheat infection in the field, another experiment was designed to determine which of the newly identified metabolites were produced *in planta*. A targeted analysis was performed on wheat grains infected with *F. graminearum* in experimental fields and wheat grains naturally infected. Additional wheat grains infected *in planta* by *F. culmorum* were also analyzed. This species is the second cause of *Fusarium* Head Blight in the world and is also a major producer of mycotoxins[2].

To the best of our knowledge, it is the first time that such an extensive analysis of the metabolome of *F. graminearum* has been achieved. The results obtained in this study open many research opportunities in terms of discovery of new drugs or toxins among the newly identified metabolites. In addition, they bring new perspectives on the impact of wheat infection by *F.*

*graminearum* and *F. culmorum*, allowing a better characterization of the possible risks for the consumers and the implementation of better control strategies.

## RESULTS AND DISCUSSION

### Untargeted Analysis of the Metabolome of *F. graminearum* in vitro

Although it is an important plant pathogen, the metabolome of *F. graminearum* has not been profusely studied. However, it presents a great interest given its virulence. Therefore, in this study, an untargeted analysis was performed in order to identify new metabolites produced by this fungus. Table 1 summarizes all the metabolites that could be identified in *F. graminearum* samples, including their retention times ( $R_T$ ), exact  $m/z$  values in  $^{12}\text{C}$  wheat substrate as well as in  $^{13}\text{C}$  wheat and  $^{13}\text{C}/^{15}\text{N}$  wheat, chemical formulas and adduct forms. Among those are included known metabolites as the mycotoxins deoxynivalenol, zearalenone, fusarin A, fusarin C and aurofusarin.

Formula determination for each metabolite was performed as previously developed and validated [23]. Briefly, after growth on  $^{12}\text{C}$  wheat grains (99%  $^{12}\text{C}$ ),  $^{13}\text{C}$  wheat grains (97%  $^{13}\text{C}$ ) or  $^{13}\text{C}/^{15}\text{N}$  wheat grains (53%  $^{13}\text{C}$ , 97 %  $^{15}\text{N}$ ), all the secondary metabolites produced by the strain of reference PH1 (INRA 605) of *F. graminearum* were either unlabeled ( $^{12}\text{C}$  cultures), fully labeled ( $^{13}\text{C}$  cultures), or doubly labeled ( $^{13}\text{C}/^{15}\text{N}$  cultures). Then, after HPLC-HRMS analysis, chromatograms of the most abundant ions in the three substrates ( $^{12}\text{C}$ ,  $^{13}\text{C}$  and  $^{13}\text{C}/^{15}\text{N}$ ) were extracted from Xcalibur (Thermo Scientific). Wheat metabolites were excluded by HPLC-HRMS analysis of a  $^{12}\text{C}$  wheat sample that was incubated without *F. graminearum*. Ions corresponding to the same metabolite in the three different substrates were grouped based on same retention time ( $R_T$ ) and similar  $\text{MS}^2$  pattern. Mass shifts could be observed among the three  $m/z$  ratios corresponding to the same metabolite. The difference between the  $^{12}\text{C}$   $m/z$  and the  $^{13}\text{C}$   $m/z$  indicated the number of atoms of carbon and the shift between the  $^{12}\text{C}$   $m/z$  and the  $^{13}\text{C}/^{15}\text{N}$   $m/z$  was used to determine the number of atoms of nitrogen. Mass Compare, an in-house-developed program (Downloadable at <http://www.toulouse.inra.fr/axiom>), was able to automatically determine the common chemical formula to the three ions, corresponding to the chemical formula of the metabolite. Finally, all the resulting chemical formulas were searched in AntiBase [24]. Additional putative annotation of metabolites was performed by interpretation of  $\text{MS}^2$  experiments, and when possible, identity was confirmed by comparison with the HPLC- $\text{MS}^2$  analysis of a standard compound.  $\text{MS}^3$  experiments were performed when further information was needed for structural determination.

Besides the known metabolites, 29 additional chemical formulas could be unambiguously associated to detected mass peaks (Table 1). The formulas of 13 of these unknown metabolites matched with formulas gathered in AntiBase (supplementary data). Since this database does not only encompass metabolites produced by filamentous fungi but also bacteria and other microorganisms, putative metabolites were limited to fungal origin and phylogenetic closeness to *Fusarium* species.

Metabolite	Formula <sup>a</sup>	Ion	$^{12}\text{C}$ $m/z$	$^{13}\text{C}$ $m/z$	$^{13}\text{C}/^{15}\text{N}$ $m/z$	Ionization mode	$R_T$ (min)	AB hits <sup>b</sup>
Known Compounds	C <sub>30</sub> H <sub>19</sub> O <sub>12</sub>	[M+H] <sup>+</sup>	571.08663	601.18773	587.13888	ESI +	32.2	1
			-0.827 ppm	-0.027 ppm	-3.286 ppm			
Aurofusarine	C <sub>30</sub> H <sub>18</sub> O <sub>12</sub> Na	[M+Na] <sup>+</sup>	593.06815	623.16843	609.12024	ESI +	32.2	
			-1.513 ppm	-2.024 ppm	-4.160 ppm			



Travail expérimental – Étude du métabolome secondaire de *F. graminearum*

Metabolite	Formula <sup>a</sup>	Ion	<sup>12</sup> C m/z	<sup>13</sup> C m/z	<sup>13</sup> C/ <sup>15</sup> N m/z	Ionization mode	R <sub>T</sub> (min)	AB hits <sup>b</sup>	
Deoxynivalenol	C <sub>15</sub> H <sub>22</sub> O <sub>6</sub>	[M+H] <sup>+</sup>	297.13326 -0.005 ppm	312.18339 -0.197 ppm	305.12924 -2.829 ppm	APCI +	12.7	1	
3-Acetyldeoxy nivalenol	C <sub>17</sub> H <sub>23</sub> O <sub>7</sub>	[M+H] <sup>+</sup>	339.1439 -1.207 ppm	356.1992 -4.6 ppm	<i>Not Detected</i>	APCI +	14.4	1	
15-Acetyldeoxy nivalenol	C <sub>17</sub> H <sub>23</sub> O <sub>7</sub>	[M+H] <sup>+</sup>	339.1439 -1.207 ppm	356.1992 -4.6 ppm	<i>Not Detected</i>	APCI +	14.2	1	
Fusarine A	C <sub>23</sub> H <sub>29</sub> O <sub>6</sub> NNa	[M+Na] <sup>+</sup>	438.18807 -1.46 ppm	461.26435 -3.293 ppm	451.22464 -3.017 ppm	ESI +	32.2	1	
Fusarine C	C <sub>23</sub> H <sub>29</sub> O <sub>7</sub> NNa	[M+Na] <sup>+</sup>	454.18316 -1.02 ppm	477.26086 0.6 ppm	467.22069 -0.483 ppm	ESI +	29.2	1	
Zearalenone	C <sub>18</sub> H <sub>21</sub> O <sub>5</sub>	[M-H] <sup>-</sup>	317.14090 2.34 ppm	335.20044 1.810 ppm	327.17363 1.941 ppm	ESI -	31.9	1	
	C <sub>36</sub> H <sub>59</sub> O <sub>7</sub> N <sub>4</sub>	[M+H] <sup>+</sup>	659.43719 -0.965 ppm	695.55713 -2.113 ppm	681.48588 -0.693 ppm	ESI +	40.4		
Fusaristatin A	C <sub>36</sub> H <sub>58</sub> O <sub>7</sub> N <sub>4</sub> Na	[M+Na] <sup>+</sup>	681.41895 -1.205 ppm	717.53839 -3.002 ppm	703.46619 -3.081 ppm	ESI +	40.4	1	
	C <sub>36</sub> H <sub>57</sub> O <sub>7</sub> N <sub>4</sub>	[M-H] <sup>-</sup>	657.42443 1.759 ppm	693.54478 1.058 ppm	680.47529 0.2 ppm	ESI -	40.4		
Newly identified compounds	C <sub>36</sub> H <sub>61</sub> O <sub>7</sub> N <sub>4</sub>	[M+H] <sup>+</sup>	661.45313 -0.524 ppm	697.57440 0.216 ppm	683.50091 -1.599 ppm	ESI +	40.6		
	Fusaristatin C	C <sub>36</sub> H <sub>60</sub> O <sub>7</sub> N <sub>4</sub> Na	[M+Na] <sup>+</sup>	683.43475 -0.982 ppm	719.55554 -0.909 ppm	705.48216 -2.533 ppm	ESI +	40.6	0
		C <sub>36</sub> H <sub>59</sub> O <sub>7</sub> N <sub>4</sub>	[M-H] <sup>-</sup>	659.43978 1.299 ppm	695.55979 0.135 ppm	682.49092 0.170 ppm	ESI -	40.6	
Fusaristatin D	C <sub>31</sub> H <sub>55</sub> O <sub>6</sub> N <sub>4</sub>	[M+H] <sup>+</sup>	579.4080 -4.7 ppm	610.5126 -4.9 ppm	<i>Not Detected</i>	ESI +	40	0	
		C <sub>31</sub> H <sub>53</sub> O <sub>6</sub> N <sub>4</sub>	[M-H] <sup>-</sup>	577.39697 -0.153 ppm	608.50021 -1.393 ppm	596.43460 -1.543 ppm	ESI -	40	
Unknown compounds	C <sub>30</sub> H <sub>52</sub> O <sub>8</sub> N <sub>6</sub> Na	[M+Na] <sup>+</sup>	647.37363 -0.392 ppm	677.47369 -1.236 ppm	668.40386 -4.361 ppm	ESI +	22.6	0	
	C <sub>18</sub> H <sub>30</sub> O <sub>4</sub> Na	[M+Na] <sup>+</sup>	333.20279 -2.523 ppm	351.26348 -1.529 ppm	343.23590 -3.725 ppm	ESI +	26.4	2	
	C <sub>18</sub> H <sub>32</sub> O <sub>4</sub> Na	[M+Na] <sup>+</sup>	335.21936 0.237 ppm	353.27914 -1.492 ppm	344.24848 -2.887 ppm	ESI +	37.58	4	
		C <sub>18</sub> H <sub>31</sub> O <sub>4</sub>	[M-H] <sup>-</sup>	311.22318 1.276 ppm	329.28354 1.126 ppm	321.25668 1.087 ppm	ESI -	37.53	
	C <sub>33</sub> H <sub>52</sub> O <sub>7</sub> N <sub>4</sub> Na	[M+Na] <sup>+</sup>	639.37268 -0.221 ppm	672.48296 -0.874 ppm	661.41954 -2.732 ppm	ESI +	37.59	0	
	C <sub>29</sub> H <sub>52</sub> O <sub>6</sub> N <sub>4</sub> Na	[M+Na] <sup>+</sup>	575.37702 -1.541 ppm	604.47401 -1.961 ppm	594.41343 -4.943 ppm	ESI +	38.52	0	
	C <sub>18</sub> H <sub>30</sub> O <sub>3</sub> Na	[M+Na] <sup>+</sup>	317.20871 -0.019 ppm	335.26868 -1.260 ppm	326.23893 0.0638 ppm	ESI +	39.54		
		C <sub>18</sub> H <sub>29</sub> O <sub>3</sub>	[M-H] <sup>-</sup>	293.21255 1.132 ppm	311.27286 0.821 ppm	303.24601 0.804 ppm	ESI -	39.54	6
		C <sub>18</sub> H <sub>29</sub> O <sub>3</sub> Na <sub>2</sub>	[M+2Na] <sup>+</sup>	339.18981 -2.508 ppm	357.24970 -3.771 ppm	349.22299 -3.490 ppm	ESI +	39.54	
		C <sub>18</sub> H <sub>29</sub> O <sub>2</sub>	[M-H <sub>2</sub> O+H] <sup>+</sup>	277.21609 -0.421 ppm	295.27618 -1.399 ppm	287.24911 -2.244 ppm	ESI +	39.54	
		C <sub>25</sub> H <sub>37</sub> O <sub>2</sub>	[M+H] <sup>+</sup>	369.27881 0.008 ppm	394.36160 -2.731 ppm	381.31725 -4.75 ppm	ESI +	40.06	2
		C <sub>38</sub> H <sub>64</sub> O <sub>7</sub> N <sub>4</sub> Na	[M+Na] <sup>+</sup>	711.46598 -1.042 ppm	749.59260 -2.140 ppm	735.51980 -2.933 ppm	ESI +	41.7	0
		C <sub>38</sub> H <sub>65</sub> O <sub>7</sub> N <sub>4</sub>	[M+H] <sup>+</sup>	689.48345 -1.924 ppm	727.21048 -2.445 ppm	713.53858 -2.007 ppm	ESI +	41.7	

Metabolite	Formula <sup>a</sup>	Ion	<sup>12</sup> C <i>m/z</i>	<sup>13</sup> C <i>m/z</i>	<sup>13</sup> C/ <sup>15</sup> N <i>m/z</i>	Ionization mode	R <sub>T</sub> (min)	AB hits <sup>b</sup>
	C <sub>28</sub> H <sub>44</sub> O <sub>4</sub> Na	[M+Na] <sup>+</sup>	467.31382 1.367 ppm	495.40688 -0.475 ppm	481.35864 -3.133 ppm	ESI +	41.7	4
	C <sub>28</sub> H <sub>42</sub> O <sub>5</sub> Na	[M+Na] <sup>+</sup>	481.29269 0.508 ppm	509.38630 -0.157 ppm	496.34192 -1.708 ppm	ESI +	43.38 43.75	3
	C <sub>21</sub> H <sub>31</sub> O <sub>2</sub>	[M-H <sub>2</sub> O+H] <sup>+</sup>	315.23209 0.740 ppm	336.30229 -0.052 ppm	326.26785 -2.788 ppm	ESI +	43.38 43.75	4
	C <sub>21</sub> H <sub>33</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	333.24171 -2.135 ppm	354.31193 -2.659 ppm	345.28213 -4.340 ppm	ESI +	43.38 43.75	
	C <sub>32</sub> H <sub>48</sub> O <sub>5</sub> Na	[M+Na] <sup>+</sup>	535.33879 -1.132ppm	567.44543 -2.325 ppm	551.39038 -1.590 ppm	ESI +	44.39	6
	C <sub>30</sub> H <sub>45</sub> O <sub>2</sub>	[M+H] <sup>+</sup>	437.34157 0.372 ppm	467.44218 0.276 ppm	453.39335 -3.824 ppm	ESI +	45.21	1
	C <sub>28</sub> H <sub>40</sub> O <sub>4</sub> Na	[M+Na] <sup>+</sup>	463.28185 -0.067 ppm	491.37485 -1.964 ppm	478.32983 -4.961 ppm	ESI +	45.72	2
	C <sub>30</sub> H <sub>45</sub> O	[M+H] <sup>+</sup>	421.34664 0.350 ppm	451.44635 -1.742 ppm	437.39829 -4.297 ppm	ESI +	47.5	0
	C <sub>28</sub> H <sub>43</sub> O <sub>2</sub>	[M+H] <sup>+</sup>	411.32536 -0.965 ppm	439.41891 -1.778 ppm	425.37241 -0.739 ppm	ESI +	47.97 50	0
	C <sub>36</sub> H <sub>60</sub> O <sub>6</sub> Na	[M+Na] <sup>+</sup>	611.42709 -1.833 ppm	647.54739 -2.461 ppm	630.49017 -2.826 ppm	ESI +	49.02 50.45	0
	C <sub>28</sub> H <sub>47</sub> O <sub>7</sub> N <sub>6</sub>	[M-H] <sup>-</sup>	579.35192 1.293 ppm	607.44541 -0.501 ppm	599.37947 -1.464 ppm	ESI -	19.3	0
	C <sub>27</sub> H <sub>45</sub> O <sub>7</sub> N <sub>6</sub>	[M-H] <sup>-</sup>	565.33607 0.971 ppm	592.42645 0.590 ppm	584.36023 -1.904 ppm	ESI -	19.45	0
	C <sub>18</sub> H <sub>21</sub> O <sub>6</sub>	[M-H] <sup>-</sup>	333.13502 1.976 ppm	351.19530 1.572 ppm	343.16755 -1.048 ppm	ESI -	21.18	5
	C <sub>18</sub> H <sub>23</sub> O <sub>6</sub>	[M-H] <sup>-</sup>	335.15060 1.755 ppm	353.21098 1.647 ppm	345.18296 -1.737 ppm	ESI -	11.89	0
	C <sub>25</sub> H <sub>43</sub> O <sub>6</sub> N <sub>6</sub>	[M-H] <sup>-</sup>	523.32552 1.077 ppm	548.40909 0.481 ppm	541.34659 -1.539 ppm	ESI -	19.2	0
	C <sub>25</sub> H <sub>35</sub> O <sub>5</sub>	[M-H] <sup>-</sup>	415.24949 1.186 ppm	440.33359 1.505 ppm	428.28264 0.007 ppm	ESI -	33.83	5
	C <sub>26</sub> H <sub>36</sub> O <sub>7</sub> Na	[M-H+HCOONa] <sup>-</sup>	483.23619 1.791 ppm	508.32020 1.978 ppm	496.27953 1.195 ppm	ESI -	33.78	
	C <sub>34</sub> H <sub>53</sub> O <sub>7</sub> N <sub>4</sub>	[M-H] <sup>-</sup>	629.39273 1.202 ppm	663.50617 0.201 ppm	651.43976 -1.135 ppm	ESI -	38.38	0
	C <sub>25</sub> H <sub>35</sub> O <sub>4</sub>	[M-H] <sup>-</sup>	399.25467 1.471 ppm	424.33848 1.242 ppm	412.29826 1.370 ppm	ESI -	39.51	0
	C <sub>18</sub> H <sub>31</sub> O <sub>3</sub>	[M-H] <sup>-</sup>	295.22827 1.361 ppm	313.28843 0.560 ppm	304.25809 0.094 ppm	ESI -	38.79	0
	C <sub>25</sub> H <sub>35</sub> O <sub>3</sub>	[M-H] <sup>-</sup>	383.25961 1.153 ppm	408.34333 0.714 ppm	396.30334 1.235 ppm	ESI -	43.19	12

**Table 1. Metabolic footprinting of *F. graminearum* on wheat *in vitro*.**

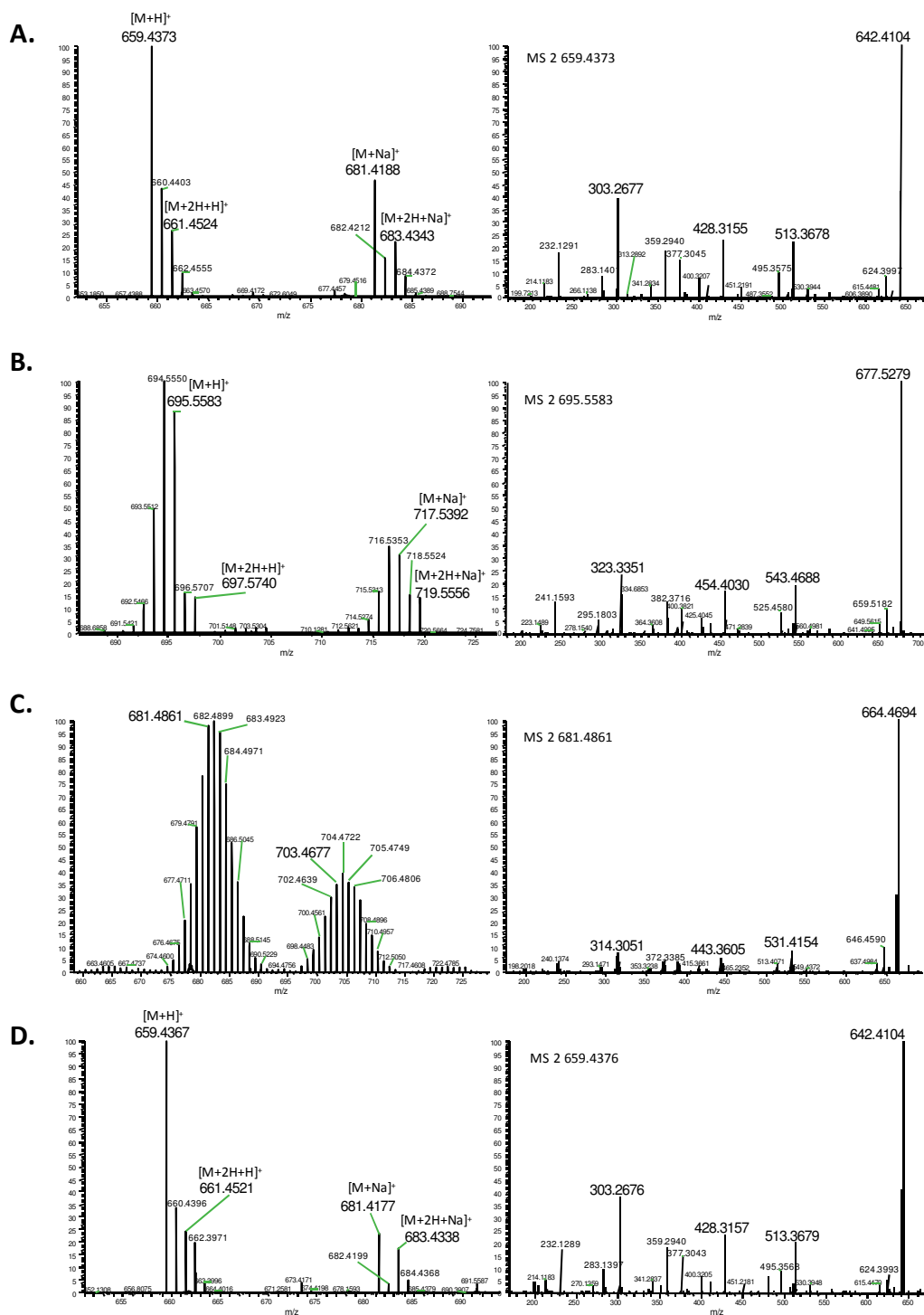
<sup>a</sup> ion formula obtained with Mass Compare.

<sup>b</sup> number of hits obtained in AntiBase for each molecule, in case there were several different adduct forms for one molecule, hits were grouped.

### Identification of fusaristatin A

Based on HPLC-HRMS results and AntiBase searches, we were able to successfully identify the presence of fusaristatin A in the metabolome of *F. graminearum* (Figure 1). This compound displayed a retention time of 40.4 min, and presented a characteristic gaussian isotopic pattern in the <sup>13</sup>C/<sup>15</sup>N sample with the major ion detected at *m/z* 681.4861. It was also detected at the same R<sub>T</sub> but at

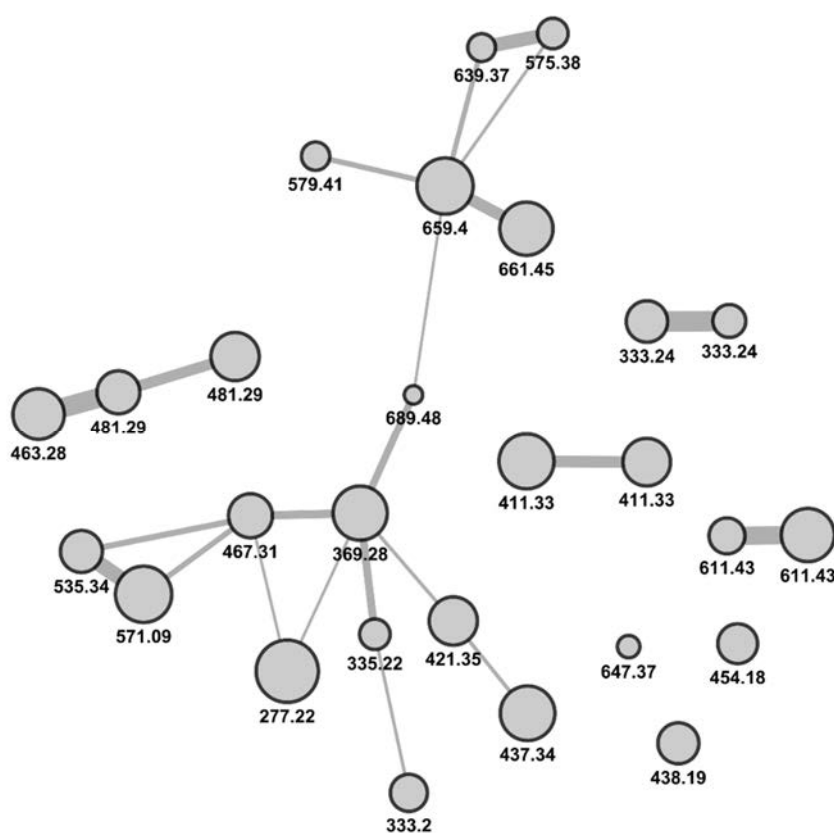
different  $m/z$  ratios in the  $^{12}\text{C}$  sample ( $m/z$  659.4373) and in the  $^{13}\text{C}$  sample ( $m/z$  695.5583). Common elemental composition of the three  $m/z$  ratios was confirmed by comparison of  $\text{MS}^2$  spectra. Additionally, this metabolite was not only ionized as the  $[\text{M}+\text{H}]^+$  form but also as the  $[\text{M}+\text{Na}]^+$  adduct ion ( $m/z$  681.4188, in the  $^{12}\text{C}$  sample). As previously explained, mass shifts between the  $^{12}\text{C}$  sample (Figure 1A) and the  $^{13}\text{C}$  sample (Figure 1B) indicated 36 atoms of carbon and the mass shift between  $^{12}\text{C}$  peaks and  $^{13}\text{C}/^{15}\text{N}$  peaks (Figure 1C) indicated four atoms of nitrogen. The chemical formula of the compound was determined using Mass Compare ( $\text{C}_{36}\text{H}_{59}\text{O}_7\text{N}_4$ ). A search request in AntiBase showed that it could correspond to the protonated form of fusaristatin A [ $\text{C}_{36}\text{H}_{58}\text{O}_7\text{N}_4+\text{H}$ ] $^+$ . The analysis of a standard of fusaristatin A in the same experimental conditions confirmed the nature of this compound (Figure 1D). Fusaristatin A was first extracted from a fungal species isolated in Germany and identified as a *Fusarium* species by the Centraalbureau voor Schimmelcultures (The Netherlands)[25]. In 2014, fusaristatin A was characterized for the first time in the secondary metabolome of *F. graminearum*[26]. Fusaristatin B ( $\text{C}_{37}\text{H}_{50}\text{O}_8\text{N}_3$ ), which was also detected in the first study, could not be identified in the present article, probably due to strain or media differences. These metabolites showed a growth-inhibitory activity in lung cancer cells (LU 65) with  $\text{IC}_{50}$  values of 23 and 7  $\mu\text{M}$ , respectively. More recently, fusaristatin A could be detected in a culture of *F. tricinctum* on rice medium[27]. The production of this metabolite was enhanced by the presence of *Bacillus subtilis* but no antibiotic properties of fusaristatin A were reported.



**Figure 1. Fusaristatin A production by *F. graminearum*.** The strain PH1 (INRA 605) of *F. graminearum* was grown at 25°C during 1 month on: (A) natural wheat grains as well as on grains labeled with stable isotopes: (B) 96.8% <sup>13</sup>C and (C) 53.4% <sup>13</sup>C and 96.8% <sup>15</sup>N. Left-hand panels represent mass spectra obtained after HPLC-HRMS analyses in positive ESI mode at 40.4 min. Right hand panels represent MS<sup>2</sup> spectra of the *m/z* ratio of interest in each substrate. The mass shift between the <sup>12</sup>C sample (659.4371) and the <sup>13</sup>C sample (695.5571) was used to determine the number of atoms of carbon (36), and the mass shift between the <sup>12</sup>C sample and the <sup>13</sup>C<sup>15</sup>N sample (681.4858) was used to determine the number of atoms of nitrogen (4). Then, combination of the three formula lists from the differently enriched substrates with Mass Compare led to the unambiguous determination of the formula (C<sub>36</sub>H<sub>60</sub>O<sub>7</sub>N<sub>4</sub>). Identification of fusaristatin A was confirmed by comparison with standard (D).

## Molecular Networking

The web-program GNPS Molecular Networking[28] allows the treatment of MS/MS spectra datasets according to the similarity of MS/MS spectra, based on the assumption that similar MS/MS spectra may come from similar molecules[29]. The network presented in Figure 2 was generated from targeted MS<sup>2</sup> experiments performed on each detected secondary metabolites. To obtain clean MS/MS fragmentation patterns (e.g. weak noise, remaining precursor ion on spectra), experiments were acquired at low resolution in the CID mode and the collision energy was optimized for each precursor ion. The fusaristatin A displayed by the node 659.4 (Figure 2) was linked to two nodes representing two unknown metabolites detected at  $m/z$  579.4 and 661.4. Despite the relatively low cosine score between these three metabolites, they displayed very similar MS/MS spectra, suggesting that they are all related to the fusaristatin family. The new secondary metabolites were putatively named fusaristatin C and fusaristatin D.

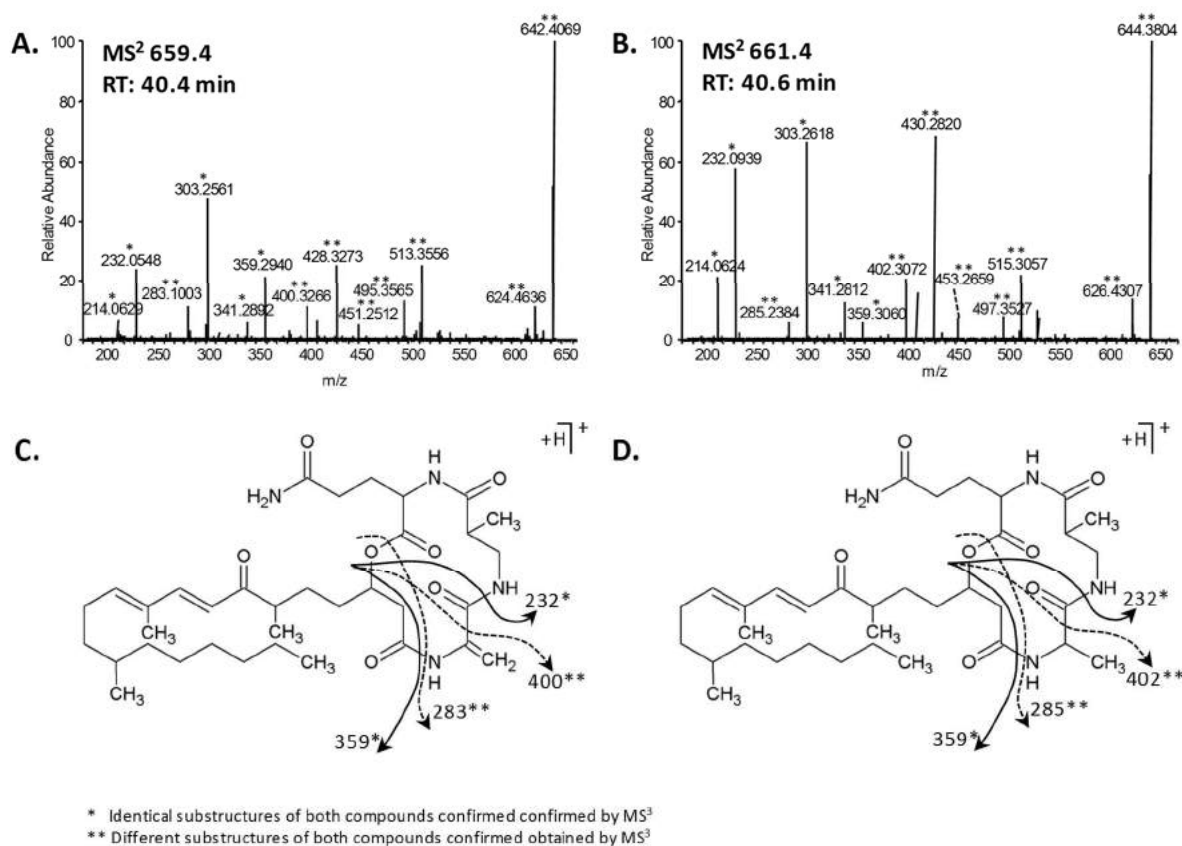


**Figure 2. Molecular network generated from secondary metabolites detected with positive electrospray.** The nodes are identified by the  $m/z$  ratio of the precursor ion. Node size represents the absolute intensity with a logarithmic scale of the ion in the HRMS spectrum. Edge width illustrates the cosine score of similarity.

## Structural analysis of the new fusaristatin C

This metabolite was systematically detected at a  $R_T$  of 40.6 min, when fusaristatin A was present. In the three substrates (<sup>12</sup>C, <sup>13</sup>C and <sup>13</sup>C/<sup>15</sup>N), it appeared with a mass shift of 2 compared to fusaristatin A, which suggested the presence of a saturated double bond (data not shown). Following the same method as before, the exact formula of this metabolite was determined to be

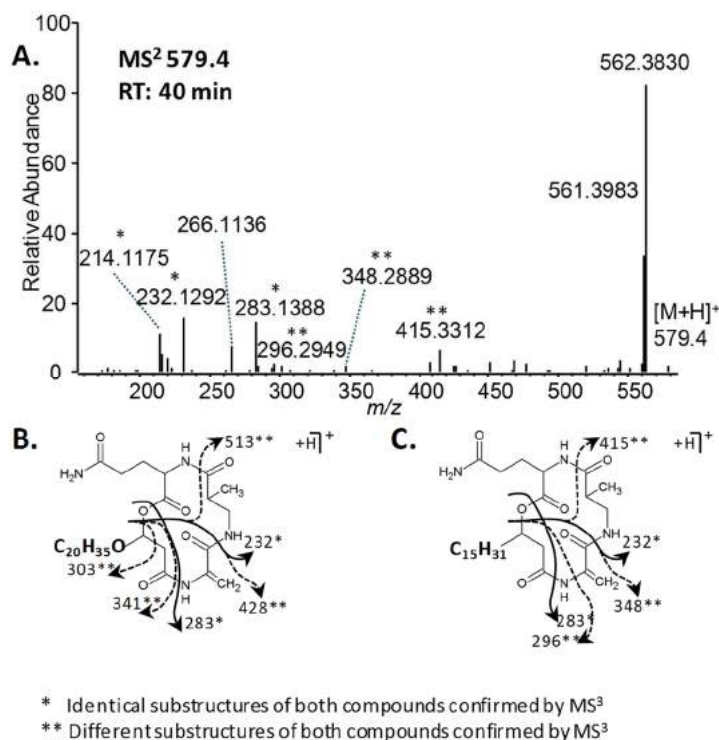
$[C_{36}H_{60}O_7N_4+H]^+$ . The MS<sup>2</sup> spectrum of this new fusaristatin (Figure 3B) presented similarities with that of fusaristatin A (Figure 3A). An MS<sup>3</sup> experiment confirmed identical substructures of the fragment ions 214, 232, 303, 341 and 359 and different substructures for the fragment ions 285, 402, 430, 435, 497, 515 and 626. Based on this, the position of the saturated double bond could be located as depicted in figure 3 (Figure 3C, D). Given these results, it was possible to elucidate the structure of fusaristatin C (Figure 3D).



**Figure 3. Structure elucidation of fusaristatin C.** The new fusaristatin was detected at 40.6 min at  $m/z$  661.4531 in the <sup>12</sup>C sample. An MS<sup>2</sup> experiment was performed to identify similarities and differences of fragmentation pattern between fusaristatin A (A) and fusaristatin C (B). As a result, and based on the structure of fusaristatin A (C), a structure for the new fusaristatin was postulated (D).

#### Structural analysis of the new Fusaristatin D

Additionally, another metabolite could be identified at a  $R_T$  of 40 min with a similar MS<sup>2</sup> fragmentation profile to that of fusaristatin A. Based on the mass shifts between  $m/z$  ratios and using Mass Compare, the chemical formula of fusaristatin D could be determined  $[C_{31}H_{54}O_6N_4+H]^+$ . Comparison of the MS<sup>2</sup> spectrum of this metabolite (Figure 4A) with that of fusaristatin A (Figure 3A), revealed the presence of similar and different fragment ions. The substructure of these fragments was analyzed by MS<sup>3</sup> which confirmed that the ions 214, 232 and 283 were identical in the two molecules, and that 296, 348 and 415 were not. Based on this, it could be deduced that the two molecules differed in the side-chain which had a formula of  $C_{20}H_{35}O$  for fusaristatin A (Figure 4B) and  $C_{15}H_{31}$  for fusaristatin D (Figure 4C). However, fragmentation of the side-chain could not be achieved given its low reactivity. Therefore, it was not possible to elucidate the complete structure of fusaristatin D.



**Figure 4 Structure elucidation of fusaristatin D.** Fusaristatin D was detected at 40 min at  $m/z$  579.4080 in the <sup>12</sup>C sample. An MS<sup>2</sup> experiment was performed (A) to identify similarities and differences of fragmentation pattern between fusaristatin A and fusaristatin D, further reflected in their structures (B and C, respectively). This comparison showed that the difference between these two metabolites was located on the side-chain. However, given the low reactivity of this chain, it was impossible to characterize the C<sub>15</sub>H<sub>31</sub> structure of fusaristatin D.

### Secondary Metabolite Production *in planta*

In order to get a better insight on the secondary metabolites produced by *F. graminearum* on wheat spikelets in the field, a second experiment was designed to determine which of the newly identified metabolites were produced *in planta*. A semi-targeted analysis was thus performed on wheat spikelets infected in experimental fields and wheat spikelets naturally infected. Naturally infected spikelets presented a co-contamination with *F. culmorum*, which is the second most important cause of the Fusarium Head Blight disease. Therefore, other wheat spikelets were also infected in experimental fields with *F. culmorum* only. Table 2 gathers all the metabolites that could be detected in the grains infected *in planta*.

Analysis of the grains infected with the strain INRA 605 of *F. graminearum* resulted in the detection of 16 metabolites, among which known mycotoxins such as deoxynivalenol, 3-acetyldeoxynivalenol, 15-acetyldeoxynivalenol and zearalenone and 12 of the newly identified metabolites. In the grains infected with the other strain INRA DGEC of *F. graminearum* isolated from the field, 26 metabolites could be detected among which aurofusarin, deoxynivalenol, 3-acetyldeoxynivalenol, 15-acetyldeoxynivalenol, zearalenone and fusarine C and 20 of the new metabolites. Given these results, it can be hypothesized that the strain INRA 605 has lost some of its pathogenicity along *in vitro* culture, which could explain why less metabolites could be detected in the grains infected with this strain compared to the strain INRA DGEC. Lower number of metabolites

detected *in planta* compared to *in vitro* may be a consequence of the interaction between the host and the pathogen. It has already been shown that, as a result of the interaction with a host plant, the profile of secondary metabolites produced by *F. graminearum* is modified[7]. These changes have been explained by the impact of plant metabolites produced during the infection process. Based on this, it is also most likely that other unknown secondary metabolites are produced when *F. graminearum* behaves as a plant pathogen. Such metabolites, host-specific and specifically related to pathogenicity, were thus not produced in the saprophytic conditions of the *in vitro* experiment and could not be characterized in this study. Therefore, further research should address the characterization of these metabolites in order to continue exploring the richness of the metabolome of *F. graminearum*.

Analysis of the grains infected with seven strains of *F. culmorum* isolated in fields resulted in the detection of 26 secondary metabolites, among which 20 of the new metabolites. The results obtained after infection with *F. culmorum* were very similar to those obtained with the strain INRA DGEC of *F. graminearum*. These observations are consistent with the fact that both *F. graminearum* and *F. culmorum* are phylogenetically close and belong to the same *Fusarium* species complex *sambucinum* section *Discolor*[18,30]. The genome of *F. culmorum* is not publicly available yet, and it would be very useful in order to further compare the secondary metabolism of both species. Given the variety of new metabolites produced by *F. culmorum* in this study, future research should address the characterization of its secondary metabolome.

Compared to the experimental infections *in planta*, less metabolites could be detected in naturally infected wheat grains. This might be partly explained by a lower production of the metabolites in natural conditions due to the presence of other competitor organisms, which may have prevented their detection with our analytical conditions. However, 50 % of the newly identified metabolites could be detected in naturally infected grains. This result brings new perspectives concerning the impact of wheat infection by *F. graminearum* and *F. culmorum* and the risks for the consumers.

Metabolite	Formula <sup>a</sup>	<sup>12</sup> C m/z	<i>F. graminearum</i>						<i>F. culmorum</i>		<i>F. graminearum</i> & <i>F. culmorum</i>				
			INRA 605 (PH1)				INRA DGEC		<i>in planta</i> n=7 %	naturally infected n=5 %					
			<i>in vitro</i>	<i>in planta</i>	n=4 <sup>b</sup>	% <sup>c</sup>	<i>in planta</i>	n=6 %							
<b>Known Compounds</b>															
Aurofusarine	C <sub>30</sub> H <sub>19</sub> O <sub>12</sub>	571.08663 -0.827 ppm	Yes	Yes	4	100	Yes	6	100	Yes	7	100	Yes	5	100
Deoxynivalenol	C <sub>15</sub> H <sub>22</sub> O <sub>6</sub>	297.13326 -0.005 ppm	Yes	Yes	4	100	Yes	6	100	Yes	7	100	Yes	2	40
3-Acetyldeoxy nivalenol	C <sub>17</sub> H <sub>23</sub> O <sub>7</sub>	339.1439 -1.207 ppm	Yes	Yes	4	100	Yes	6	100	Yes	7	100	Yes	1	20
15-Acetyldeoxy nivalenol	C <sub>17</sub> H <sub>23</sub> O <sub>7</sub>	339.1439 -1.207 ppm	Yes	Yes	4	100	Yes	6	100	No	0	0	No	0	0
Fusarine A	C <sub>23</sub> H <sub>29</sub> O <sub>6</sub> NNa	438.18807 -1.46 ppm	Yes	No	0	0	No	0	0	Yes	2	29	No	0	0
Fusarine C	C <sub>23</sub> H <sub>29</sub> O <sub>7</sub> NNa	454.18316 -1.02 ppm	Yes	No	0	0	Yes	2	33	Yes	7	100	No	0	0
Zearalenone	C <sub>18</sub> H <sub>21</sub> O <sub>5</sub>	317.14090 2.34 ppm	Yes	Yes	4	100	Yes	5	83	Yes	7	100	Yes	3	60
Fusaristatin A	C <sub>36</sub> H <sub>59</sub> O <sub>7</sub> N <sub>4</sub>	659.43719 -0.965 ppm	Yes	Yes	4	100	Yes	6	100	Yes	7	100	Yes	5	100



**Travail expérimental – Étude du métabolome secondaire de *F. graminearum***

Metabolite	Formula <sup>a</sup>	<sup>12</sup> C m/z	<i>F. graminearum</i>						<i>F. culmorum</i>			<i>F. graminearum</i> & <i>F. culmorum</i>				
			INRA 605 (PH1)				INRA DGEC		<i>in planta</i> n=7 %			naturally infected n=5 %				
			<i>in vitro</i>	<i>in planta</i>	n=4 <sup>b</sup>	% <sup>c</sup>	<i>in planta</i>	n=6	%	<i>in planta</i>	n=7	%	naturally infected	n=5	%	
<b>Newly identified compounds</b>	Fusaristatin C	C <sub>36</sub> H <sub>61</sub> O <sub>7</sub> N <sub>4</sub>	661.45313 -0.524 ppm	Yes	Yes	4	100	Yes	6	100	Yes	7	100	Yes	5	100
	Fusaristatin D	C <sub>31</sub> H <sub>53</sub> O <sub>6</sub> N <sub>4</sub>	577.39697 -0.153 ppm	Yes	No	0	0	Yes	5	83	Yes	6	86	Yes	1	20
<b>Unknown compounds</b>		C <sub>30</sub> H <sub>52</sub> O <sub>8</sub> N <sub>6</sub> Na	647.37363 -0.392 ppm	Yes	Yes	4	100	Yes	6	100	Yes	7	100	Yes	5	100
		C <sub>18</sub> H <sub>30</sub> O <sub>4</sub> Na	333.20279 -2.523 ppm	Yes	No	0	0	No	0	0	No	0	0	No	0	0
		C <sub>18</sub> H <sub>31</sub> O <sub>4</sub>	311.22318 1.276 ppm	Yes	No	0	0	Yes	6	100	Yes	7	100	No	0	0
		C <sub>33</sub> H <sub>52</sub> O <sub>7</sub> N <sub>4</sub> Na	639.37268 -0.221 ppm	Yes	No	0	0	No	0	0	No	0	0	No	0	0
		C <sub>29</sub> H <sub>52</sub> O <sub>6</sub> N <sub>4</sub> Na	575.37702 -1.541 ppm	Yes	No	0	0	Yes	3	50	Yes	6	86	No	0	0
		C <sub>18</sub> H <sub>29</sub> O <sub>3</sub>	293.21255 1.132 ppm	Yes	Yes	4	100	Yes	6	100	Yes	7	100	Yes	5	100
		C <sub>25</sub> H <sub>37</sub> O <sub>2</sub>	369.27881 0.008 ppm	Yes	Yes	4	100	Yes	6	100	Yes	7	100	Yes	5	100
		C <sub>38</sub> H <sub>65</sub> O <sub>7</sub> N <sub>4</sub>	689.48345 -1.924 ppm	Yes	No	0	0	Yes	3	50	Yes	5	71	No	0	0
		C <sub>28</sub> H <sub>44</sub> O <sub>4</sub> Na	467.31382 1.367 ppm	Yes	No	0	0	No	0	0	No	0	0	No	0	0
		C <sub>28</sub> H <sub>42</sub> O <sub>5</sub> Na	481.29269 0.508 ppm	Yes	Yes	4	100	Yes	6	100	Yes	7	100	No	0	0
		C <sub>21</sub> H <sub>33</sub> O <sub>3</sub>	333.24171 -2.135 ppm	Yes	No	0	0	No	0	0	No	0	0	No	0	0
		C <sub>32</sub> H <sub>48</sub> O <sub>5</sub> Na	535.33879 -1.132 ppm	Yes	No	0	0	No	0	0	No	0	0	No	0	0
		C <sub>30</sub> H <sub>45</sub> O <sub>2</sub>	437.34157 0.372 ppm	Yes	No	0	0	Yes	5	83	Yes	6	86	No	0	0
		C <sub>28</sub> H <sub>40</sub> O <sub>4</sub> Na	463.28185 -0.067 ppm	Yes	Yes	4	100	Yes	6	100	Yes	7	100	Yes	1	20
		C <sub>30</sub> H <sub>45</sub> O	421.34664 0.350 ppm	Yes	No	0	0	Yes	4	67	Yes	6	86	No	0	0
		C <sub>28</sub> H <sub>43</sub> O <sub>2</sub>	411.32536 -0.965 ppm	Yes	Yes	4	100	Yes	6	100	Yes	7	100	Yes	5	100
		C <sub>36</sub> H <sub>60</sub> O <sub>6</sub> Na	611.42709 -1.833 ppm	Yes	No	0	0	Yes	6	100	No	0	0	No	0	0
		C <sub>28</sub> H <sub>47</sub> O <sub>7</sub> N <sub>6</sub>	579.35192 1.293 ppm	Yes	Yes	4	100	Yes	6	100	Yes	7	100	Yes	5	100
		C <sub>27</sub> H <sub>45</sub> O <sub>7</sub> N <sub>6</sub>	565.33607 0.971 ppm	Yes	Yes	4	100	Yes	6	100	Yes	7	100	No	0	0
		C <sub>18</sub> H <sub>21</sub> O <sub>6</sub>	333.13502 1.976 ppm	Yes	No	0	0	No	0	0	No	0	0	No	0	0
	C <sub>18</sub> H <sub>23</sub> O <sub>6</sub>	335.15060 1.755 ppm	Yes	No	0	0	No	0	0	No	0	0	No	0	0	
	C <sub>25</sub> H <sub>43</sub> O <sub>6</sub> N <sub>6</sub>	523.32552 1.077 ppm	Yes	Yes	3	75	Yes	6	100	Yes	7	100	Yes	5	100	
	C <sub>25</sub> H <sub>35</sub> O <sub>5</sub>	415.24949 1.186 ppm	Yes	No	0	0	No	0	0	No	0	0	No	0	0	

Metabolite	Formula <sup>a</sup>	<sup>12</sup> C m/z	<i>F. graminearum</i>						<i>F. culmorum</i>		<i>F. graminearum</i> & <i>F. culmorum</i>				
			INRA 605 (PH1)				INRA DGEC		<i>in planta</i>	n=7	%	naturally infected	n=5	%	
			<i>in vitro</i>	<i>in planta</i>	n=4 <sup>b</sup>	% <sup>c</sup>	<i>in planta</i>	n=6							%
	C <sub>34</sub> H <sub>53</sub> O <sub>7</sub> N <sub>4</sub>	629.39273 1.202 ppm	Yes	No	0	0	Yes	4	67	Yes	7	100	No	0	0
	C <sub>25</sub> H <sub>35</sub> O <sub>4</sub>	399.25467 1.471 ppm	Yes	No	0	0	No	0	0	Yes	7	100	No	0	0
	C <sub>18</sub> H <sub>31</sub> O <sub>3</sub>	295.22827 1.361 ppm	Yes	Yes	4	100	Yes	6	100	Yes	7	100	Yes	5	100
	C <sub>25</sub> H <sub>35</sub> O <sub>3</sub>	383.25961 1.153 ppm	Yes	No	0	0	No	0	0	No	0	0	No	0	0

**Table 2. Metabolic footprinting of *F. graminearum* and *F. culmorum* on wheat *in planta*.**

<sup>a</sup> ion formula of the indicated ion obtained with Mass Compare.

<sup>b</sup> number of analyzed samples for each experimental condition.

<sup>c</sup> percentage of analyzed samples in which each metabolite could be detected.

## CONCLUSION

This study illustrated the interest of exploring the secondary metabolome of economically important and challenging mycotoxinogenic phytopathogens such as *F. graminearum* and *F. culmorum*[8,31,32]. The combination of high accuracy mass spectrometry and double isotopic labeling efficiently enabled the identification of fusaristatin A as well as 29 new metabolites, and the characterization of the new fusaristatin C and D. There is very limited knowledge on the toxicity of fusaristatin A and none for the two other metabolites. However, given their structural similarities, it is possible to suggest that the toxicity of the fusaristatins C and D is related to that of fusaristatin A that is, inhibition of lung cells proliferation. More knowledge is needed concerning the toxicity of fusaristatins as they may represent a risk after spore inhalation or after ingestion of contaminated food/feed, especially since these metabolites could be detected in wheat grains infected *in planta* and naturally infected grains. Further analyses revealed that more than 50% of the new secondary metabolites are produced *in planta* while routine analyses of food and feed only target a limited number of mycotoxins produced by *F. graminearum* and *F. culmorum*[8,33,34]. The presence of these new metabolites is not to be ignored since they could have toxic properties of their own or could interact with the toxicity of known mycotoxins. Therefore, more knowledge is needed concerning the toxicity of the new metabolites in order to implement, if necessary, new control measures for consumers' protection.

In a general manner, given the impact of *F. graminearum* and *F. culmorum*, the results obtained in this study should help guide future research on the secondary metabolites produced by these two species. The application of the method presented here also offers great opportunities for the characterization of other challenging fungi.

## MATERIALS AND METHODS

### Chemicals and Reagents

All solvents and chemicals were purchased from Fisher Scientific (Thermo Fisher Scientific, Illkirch, France). Solvents were HPLC grade except acetonitrile, which was MS grade. All other chemicals were analytical grade.

### ***In vitro* cultures of *F. graminearum* on $^{12}\text{C}$ , $^{13}\text{C}$ and $^{13}\text{C}/^{15}\text{N}$ labeled Wheat Grains**

The strain of *Fusarium graminearum* used in this study (NRRL 31084, designated as INRA 605) was provided by the Centraalbureau voor Schimmelcultures (Utrecht, The Netherlands). Seven-days cultures of INRA 605 on Potato Dextrose Agar (PDA) (Sigma-Aldrich, St Quentin-Fallavier, France) were used to inoculate three different wheat grains substrates (variety Caphorn of *Triticum aestivum*): (1) naturally enriched wheat grains (99%  $^{12}\text{C}$ / 1%  $^{13}\text{C}$ , referred to as  $^{12}\text{C}$  wheat), (2) nearly 100 %  $^{13}\text{C}$  enriched wheat grains (referred to as  $^{13}\text{C}$  wheat) and (3) nearly 50 %  $^{13}\text{C}$  and nearly 100 %  $^{15}\text{N}$  wheat grains (referred to as  $^{13}\text{C}/^{15}\text{N}$  wheat). Wheat uniform isotopic enrichments were performed in hermetically sealed chambers during a complete growth cycle with labeled  $\text{CO}_2$  and labeled nitrate and ammonium salts as previously described[23]. Final isotopic enrichments were measured with a Delta V Advantage isotope ratio mass spectrometer (Thermo Fisher Scientific). They were of  $96.8\% \pm 0.15$  for the  $^{13}\text{C}$  wheat,  $53.4\% \pm 0.3$   $^{13}\text{C}$  and  $96.8\% \pm 0.07$   $^{15}\text{N}$  for the  $^{13}\text{C}/^{15}\text{N}$  wheat. Grain preparation and inoculations were performed as previously described[23]. Prior to inoculation, wheat grains were washed, soaked in distilled water and autoclaved twice to get rid of any endogenous living microorganisms. Water activity ( $a_w$ ) of the grains was maintained above 0.98 throughout the experiment. Cultures were allowed to grow 1 month at  $25^\circ\text{C}$  in the dark without shaking. In order to discriminate between metabolites produced by *F. graminearum* and wheat metabolites,  $^{12}\text{C}$  wheat grains were also incubated in the same conditions as the other samples but without fungi.

### **Secondary Metabolites Production on Wheat Spikelets Infected *in planta***

Infected wheat spikelets used in this experiment originated from different field studies: (1) wheat spikelets infected with the strain INRA 605, (2) wheat spikelets infected with a naturally occurring strain of *F. graminearum*, (3) wheat spikelets infected with a strain of *F. culmorum*, (4) wheat spikelets naturally infected with *F. graminearum* and *F. culmorum*.

For experiments 1, 2 and 3 intact spikelets were inoculated at anthesis with 1 ml spray of conidia suspension containing  $10^5$  conidia/ml.

### **Metabolite Extraction from *in vitro* cultures and wheat spikelets**

Metabolite extraction from *in vitro* cultures was performed, as previously described[35], in a fume hood using 200 ml of chloroform at room temperature in a 1l glass bottle where the blend of the mycelia and the grains was inserted cut in small pieces. The bottle was placed on a shaking table IKA Labortechnik HS501 (IKA, Staufen, Germany) at 180 rpm for 36 h. The chloroform extract was then filtered and evaporated in two steps: first with a rotavapor R-210/215 (Buchi, Flawil, Switzerland) and then with a Zymark TurboVap-LV Evaporator (McKinley Scientific, Sparta, NJ, USA). The resulting dry residue (metabolites) was dissolved in 400  $\mu\text{l}$  of methanol and then filtered with 0.45  $\mu\text{m}$  Whatman filters (Whatman, GE Healthcare, Kent, UK) before HPLC-HRMS analysis.

Metabolite extraction from wheat spikelets infected *in planta* was performed on 2 g of grains from these wheat spikes. Grains were milled with a Retsch MM400 mixer mill (Retsch France, Eragny sur Oise, France) for 4 min at 25 Hz. Then, extraction was performed with 30 ml of chloroform as described above for the *in vitro* experiment.

### **Identification of known compounds and preparation of metabolite reference compounds**

AntiBase 2012, which is one of the most exhaustive databases for metabolites produced by microorganisms[24], as well as articles about the secondary metabolome of *F. graminearum* were used for annotation of the metabolites. Metabolite reference compounds have been collected over the

years either from commercial sources, such as Sigma (Sigma-Aldrich, Saint-Quentin Fallavier, France) or as gifts from different research groups. Fusaristatin A was a gift from Shiono, Y. from the Department of Bioresource Engineering, Faculty of Agriculture, Yamagata University, Japan. All standards were dissolved in methanol for qualitative analysis by HPLC-MS as described below. Standard solutions were kept at -20°C.

#### **Untargeted HPLC and mass spectrometry analysis of *in vitro* cultures**

Untargeted analyses were performed using high performance liquid chromatography (HPLC) coupled to an LTQ Orbitrap XL high-resolution mass spectrometer (HRMS) (Thermo Fisher Scientific, Les Ulis, France). Briefly, 10 µL of methanol suspension of culture extracts were injected into a reversed phase (150\*2.0 mm) 5 µm Luna C18 column (Phenomenex, Torrance, CA, USA) operated at a flow rate of 0.2 mL/min. A gradient program was performed with 0.1 % formic acid (phase A) and 100 % acetonitrile (phase B) with the following elution gradient: 0 min 20% B, 30 min 50% B, between 35-45 min 90% B, between 50-60 min 20% B. HRMS acquisitions were achieved with electrospray ionization (ESI) in the positive and negative modes as follows: spray voltage +4.5 kV, capillary temperature 350°C, sheath gas (N<sub>2</sub>) flow rate 40 au (arbitrary units), auxiliary gas (N<sub>2</sub>) flow rate 6 au in the positive mode, and spray voltage -3.7 kV, capillary temperature 350°C, sheath gas (N<sub>2</sub>) flow rate 30 au, auxiliary gas (N<sub>2</sub>) flow rate 10 au in the negative mode. Full MS spectra were acquired at a resolution of 60000 with a range of mass-to-charge ratio (*m/z*) set to 50 – 800. The high-resolution mass analyzer was calibrated in each ionization mode based on the supplier's protocol and calibration mixtures (Thermo Fisher Scientific, Les Ulis, France).

For the detection of deoxynivalenol and its acetylated derivatives, a reversed phase (150\*2.0 mm) 5 µm polar Synergy column (Phenomenex, Torrance, CA, United States) was used instead of the Luna C18 column for being more efficient to separate polar compounds. A gradient program was performed with 0.1% acetic acid (phase A) and acetonitrile (phase B) as mobile phases. The elution started with an isocratic elution 100% of solvent A for 5 min. Then solvent B was increased linearly to 50% within 5 min. After 6 min at 50%, the gradient was increased again linearly to 90% of B within 4 min. Finally, after 5 min of 90% solvent B, the gradient was decreased to initial conditions in 2 min and remained at this value for the last 10 min. Atmospheric pressure chemical ionization (APCI) operating in the positive mode was used to perform HRMS analyses. Ionization conditions were: 260°C for the vaporizer temperature, 2.5 µA for the corona discharge, 32 au for the sheath gas flow rate, 5 au for the auxiliary gas and 290°C for the capillary temperature. Mass spectrometric resolution was set to 60,000 for *m/z* range from 50 to 800.

#### **Targeted HPLC and mass spectrometry analysis of *in planta* samples**

A targeted analysis method was developed to analyze the metabolites produced by *F. graminearum in planta* according to the list of metabolites detected in the *in vitro* experiment. The analysis was performed using high performance liquid chromatography (HPLC) coupled to a TSQ Vantage Triple Stage Quadrupole mass spectrometer (Thermo Fisher Scientific) by Multiple Reaction Monitoring (MRM). Metabolites identification was based on retention time and detection of the three most abundant fragment ions. Parent-parent transitions were used when less than three fragment ions were detected. Collision energies were optimized for each fragmentation ion. Chromatographic and mass spectrometric conditions remained identical to the ones used in the untargeted method. For this experiment, a Triple Stage Quadrupole was chosen instead of the LTQ Orbitrap due to its higher sensitivity necessary to detect the metabolites that were less abundant than in the *in vitro* experiment.

## Molecular Networking

The GNPS parameters were optimized as follow: parent mass tolerance = 0.002 Da; fragment ion mass tolerance = 0.9, minimum cosine score = 0.6; minimum match fragment ion = 2; maximum connected components = 100; Top K = 100 and Filter Peaks in 50Da Window was removed. Visualization of molecular networks was performed with Cytoscape 2.8.2[36].

## ACKNOWLEDGEMENTS

Patricia Cano was supported by doctoral fellowships co-funded by the Agence Nationale de Sécurité Sanitaire de l'Alimentation, de l'Environnement et du Travail (ANSES) and the Institut National de la Recherche Agronomique (INRA). This work was supported by grants (ES-2007-063 and EST-2011/1/087) from ANSES. HPLC-HRMS analyses were performed within the MetaToul-AXIOM analytical platform of French National Infrastructure of Metabolomics and Fluxomics (MetaboHUB-ANR-11-INBS-0010). We also would like to thank Marisa Younanian from the University of Auburn (Auburn, Alabama, USA) for her help correcting the manuscript.

## COMPETING INTERESTS

The authors declare no competing financial interests.

## REFERENCES

1. CAST (2003) in *Potential economic costs of mycotoxins in United States* (Council for Agricultural Science and Technology), pp. 136-142.
2. Leslie JF & Summerell BA (2006) *The Fusarium laboratory manual* (Blackwell Publishing, Oxford, UK).
3. Goswami RS & Kistler HC (2004) Heading for disaster: *Fusarium graminearum* on cereal crops. *Molecular Plant Pathology* 5(issue): 515-525.
4. Dean R, Van Kan JA, Pretorius ZA, Hammond-Kosack KE, Di Pietro A, Spanu PD, Rudd JJ, Dickman M, Kahmann R, Ellis J, *et al.* (2012) The Top 10 fungal pathogens in molecular plant pathology. *Mol Plant Pathol* 13(issue): 414-430.
5. Leslie JF & Summerell BA (2013) An overview of *Fusarium*. *Fusarium: Genomics, Molecular and Cellular Biology*, eds. Brown DW & Proctor RH (Caister Academic Press, Norfolk, UK), pp. 1-10.
6. Urban M & Hammond-Kosack KE (2013) Molecular genetics and genomic approaches to explore *Fusarium* infection of wheat floral tissue. *Fusarium: Genomics, Molecular and Cellular Biology*, eds. Brown DW & Proctor RH (Caister Academic Press, Norfolk, UK), pp. 43-80.
7. Kazan K, Gardiner DM, & Manners JM (2011) On the trail of a cereal killer: recent advances in *Fusarium graminearum* pathogenomics and host resistance. *Molecular Plant Pathology* 13(issue): 399-413.
8. SCOOP (2003) Collection of occurrence data of *Fusarium* toxins in food and assessment of dietary intake by the population of EU member states. *Toxicol Lett* 153(issue): 133-143.
9. Desjardins AE & Proctor RH (2007) Molecular biology of *Fusarium* mycotoxins. *Int. J. Food Microbiol.* 119(issue): 47-50.
10. Zinedine A, Soriano JM, Molto JC, & Manes J (2007) Review on the toxicity, occurrence, metabolism, detoxification, regulations and intake of zearalenone: An oestrogenic mycotoxin. *Food Chem Toxicol* 45(issue): 1-18.
11. Maresca M (2013) From the gut to the brain: journey and pathophysiological effects of the food-associated trichothecene mycotoxin deoxynivalenol. *Toxins (Basel)* 5(issue): 784-820.
12. IARC (1993) Toxins derived from *Fusarium moniliforme*: fumonisins B1 and B2 and fusarin C. *IARC Monogr Eval Carcinog Risks Hum* 56(issue): 445-466.
13. Pestka JJ & Smolinski AT (2005) Deoxynivalenol: toxicology and potential effects on humans. *J Toxicol Environ Health B Crit Rev* 8(issue): 39-69.
14. Trail F (2009) For blighted waves of grain: *Fusarium graminearum* in the postgenomics era. *Plant Physiol* 149(issue): 103 - 110.
15. Boenisch M & Schafer W (2011) *Fusarium graminearum* forms mycotoxin producing infection structures on wheat. *BMC Plant Biology* 11(issue): 110.
16. Arseniuk E, Foremska E, G T, Óral, & Chełkowski J (1999) *Fusarium* Head Blight Reactions and Accumulation of Deoxynivalenol (DON) and Some of its Derivatives in Kernels of Wheat, Triticale and Rye. *Journal of Phytopathology* 147(issue): 577-590.

17. Ma L-J, Geiser DM, Proctor RH, Rooney AP, O'Donnell K, Trail F, Gardiner DM, Manners JM, & Kazan K (2013) *Fusarium* Pathogenomics. *Annual Review of Microbiology* 67(issue): 399-416.
18. Eriksen GS & Alexander NJ (1998) *Fusarium Toxins in Cereals: A Risk Assessment* (Nordic Council of Ministers, Copenhagen, Denmark).
19. Cuomo CA, Gueldener U, Xu JR, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma LJ, Baker SE, Rep M, et al. (2007) The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. *Science* 317(issue): 1400-1402.
20. Hansen FT, Sorensen JL, Giese H, Sondergaard TE, & Frandsen RJN (2012) Quick guide to polyketide synthase and nonribosomal synthetase genes in *Fusarium*. *Int. J. Food Microbiol.* 155(issue): 128-136.
21. Gaffoor I, Brown DW, Plattner R, Proctor RH, Qi WH, & Trail F (2005) Functional analysis of the polyketide synthase genes in the filamentous fungus *Gibberella zeae* (Anamorph *Fusarium graminearum*). *Eukaryot. Cell* 4(issue): 1926-1933.
22. Sørensen JL, Hansen FT, Sondergaard TE, Staerk D, Lee TV, Wimmer R, Klitgaard LG, Purup S, Giese H, & Frandsen RJN (2012) Production of novel fusarielins by ectopic activation of the polyketide synthase 9 cluster in *Fusarium graminearum*. *Environmental Microbiology* 14(issue): 1159-1170.
23. Cano PM, Jamin EL, Tadrist S, Bourdaud'hui P, Pean M, Debrauwer L, Oswald IP, Delaforge M, & Puel O (2013) New Untargeted Metabolic Profiling Combining Mass Spectrometry and Isotopic Labeling: Application on *Aspergillus fumigatus* Grown on Wheat. *Anal Chem* 85(issue): 8412-8420.
24. Laatsch H (2012) (Wiley-VCH, Weinheim, Germany; <http://www.wiley-vch.de/stmdata/antibase.php>).
25. Shiono Y, Tsuchinari M, Shimanuki K, Miyajima T, Murayama T, Koseki T, Laatsch H, Funakoshi T, Takanami K, & Suzuki K (2007) Fusaristatins A and B, Two New Cyclic Lipopeptides from an Endophytic *Fusarium sp.* *J Antibiot* 60(issue): 309-316.
26. Ola ARB, Thomy D, Lai D, Brötz-Oesterhelt H, & Proksch P (2013) Inducing Secondary Metabolite Production by the Endophytic Fungus *Fusarium tricinctum* through Coculture with *Bacillus subtilis*. *Journal of Natural Products* 76(issue): 2094-2099.
27. O'Donnell K, Rooney AP, Proctor RH, Brown DW, McCormick SP, Ward TJ, Frandsen RJN, Lysoe E, Rehner SA, Aoki T, et al. (2013) Phylogenetic analyses of RPB1 and RPB2 support a middle Cretaceous origin for a clade comprising all agriculturally and medically important *fusaria*. *Fungal Genet. Biol.* 52(issue): 20-31.
28. Streit E, Schatzmayr G, Tassis P, Tzika E, Marin D, Taranu I, Tabuc C, Nicolau A, Aprodu I, Puel O, et al. (2012) Current situation of mycotoxin contamination and co-occurrence in animal feed--focus on Europe. *Toxins (Basel)* 4(issue): 788-809.
29. Streit E, Schwab C, Sulyok M, Naehrer K, Krska R, & Schatzmayr G (2013) Multi-Mycotoxin Screening Reveals the Occurrence of 139 Different Secondary Metabolites in Feed and Feed Ingredients. *Toxins* 5(issue): 504-523.
30. Marin S, Ramos AJ, Cano-Sancho G, & Sanchis V (2013) Mycotoxins: Occurrence, toxicology, and exposure assessment. *Food and Chemical Toxicology* (issue).
31. Zachariasova M, Lacina O, Malachova A, Kostelanska M, Poustka J, Godula M, & Hajslova J (2010) Novel approaches in analysis of *Fusarium* mycotoxins in cereals employing ultra performance liquid chromatography coupled with high resolution mass spectrometry. *Analytica Chimica Acta* 662(issue): 51-61.
32. Gauthier T, Wang X, Sifuentes Dos Santos J, Fysikopoulos A, Tadrist S, Canlet Cc, Artigot MP, Loiseau N, Oswald IP, & Puel O (2012) Trypacidin, a Spore-Borne Toxin from *Aspergillus fumigatus*, Is Cytotoxic to Lung Cells. *PLoS ONE* 7(issue): e29906.

## SUPPLEMENTARY DATA

	Formula	<sup>12</sup> C m/z	AB hits	Putative metabolite	Producing Organism	Reference
ESI +	C <sub>18</sub> H <sub>30</sub> O <sub>4</sub> Na	333.20279	2	Hydrophorone C12, Hydrophorone D12	<i>Hygrophorus pustulatus</i> (Basidiomycete)	Lübken, T. et al., <i>Phytochemistry</i> 65 (2004) 1061-71; Lübken, T. et al., <i>J. Mass Spectrom.</i> 41 (2006) 361-371
				(14RS)-(10E,12E)-14-Hydroxy-9-oxo-10,12-octadecadienoic acid	<i>Pleurocybella porrigens</i>	Amakura, Y. et al., <i>Chem. Pharm. Bull.</i> 54 (2006) 1213-1215
	C <sub>18</sub> H <sub>32</sub> O <sub>4</sub> Na	335.21936	4	Filoboletic acid; (9Z,11E)-8,13-dihydroxyoctadeca-9,11-dienoic acid	<i>Filoboletus sp.</i>	Simon, B. et al., <i>Phytochemistry</i> 36 (1994) 815-816
				Piliferolide C	<i>Ophiostoma piliferum</i>	Ayer, W.A. et al., <i>Heterocycles</i> 39 (1994) 561-9
				Hydrophorone A12	<i>Hygrophorus persoonii</i> (Basidiomycete)	Lübken, T. et al., <i>Phytochemistry</i> 65 (2004) 1061-71; Lübken, T. et al., <i>J. Mass Spectrom.</i> 41 (2006) 361-371

Travail expérimental – Étude du métabolome secondaire de *F. graminearum*

Formula	<sup>12</sup> C m/z	AB hits	Putative metabolite	Producing Organism	Reference
<b>C<sub>18</sub>H<sub>30</sub>O<sub>3</sub>Na</b>	317.20871	6	6-Hydroxy-4-oxo-11-octadecenoic acid	<i>Hygrophorus discoxanthus</i> (Basidiomycete)	Gilardoni, G. et al., Nat. Prod. Comm. 1 (2006) 1079-1084
			Conocandin; A-32287	<i>Hormococcus conorum</i>	Banfi, L. et al., J. Org. Chem. 1987, 52, 5452-7
			(1S,2S)-3-Oxo-2(2Z-pentenyl)cyclopentane-1-octanoic acid	<i>Fusarium oxysporum</i>	Miersch, O. et al., Phytochem. 50 (1999) 517-23
			11-Keto-9(E),12(E)-octadecadienoic acid	<i>Trichoderma</i>	Shinohara, C. et al., J. Antibiot. 52 (1999) 171-4; Chikanishi, T. et al., J. Antibiot. 52 (1999) 797-802
			9-Oxo-10,12-octadecadienoic acid	<i>Clitocybe clavipes</i>	Kawagishi, H. et al., J. Nat. Prod. 65 (2002) 1712-1714
			Hygrophorone F12	<i>Hygrophorus persoonii</i> (Basidiomycete)	Lübken, T. et al., Phytochemistry 65 (2004) 1061-71; Lübken, T. et al., J. Mass Spectrom. 41 (2006) 361-371
<b>C<sub>25</sub>H<sub>37</sub>O<sub>2</sub></b>	369.27881	2	13-Oxo-9(Z),11(E)-octadecadienoic acid	<i>Ripartites metrodii</i> , <i>R. tricholoma</i>	Weber, D. et al., Z. Naturf. 61C (2006) 663-669
			Variocolin; EA-1	<i>Aspergillus varicolor</i> ; <i>Emericella aurantio-brunnea</i>	Hensens, O. et al., J. Org. Chem. (1991), 56, 3399; Hensens, O. et al., J. Org. Chem. 56 (1991) 3399-3403; Fujimoto, H. et al., Chem. Pharm. Bull. 48 (2000) 1436-41
			Variocolin	<i>Emericella aurantio-brunnea</i>	Fujimoto, H. et al., Tennen Yuki Kagobutsu Toronkai Koen Yoshishu 42 (2000) 481-6
<b>C<sub>28</sub>H<sub>44</sub>O<sub>4</sub>Na</b>	467.31382	4	Mer-NF-8054-A	<i>Aspergillus ustus</i>	Sakai, K. et al., J. Antibiot. 47 (1994) 591-594; Sakamoto, M. et al., Jpn. Kokai Tokkyo Koho JP (1994) 06 145,194
			3beta,5a,9a-Trihydroxy-(22E)-ergosta-7,22-diene-6-one	<i>Hericium erinaceum</i>	Yaoita, Y. et al., Chem. Pharm. Bull. 50 (2002) 551-3
			5a,6a;8a,9a-Diepoxy-(22E,24R)-ergost-22-ene-3beta,7a-diol	<i>Amanita pantherina</i> , <i>Amanita virgineoides</i> , <i>Lactarius piperatus</i> , <i>Lyophyllum</i>	Yaoita, Y. et al., Chem. Pharm. Bull. 47 (1999) 847-851
			5a,9a-Epidioxy-(22E)-ergosta-7,22-diene-3beta,6beta-diol	<i>Pleurotus eryngii</i>	Yaoita, Y. et al., Chem. Pharm. Bull. 50 (2002) 551-3
<b>C<sub>28</sub>H<sub>42</sub>O<sub>5</sub>Na</b>	481.29269	3	Ergokonin B	<i>Trichoderma koningii</i>	Augustiniak, H. et al., Liebigs Ann. Chem. 4 (1991) 361-6
			Mer-NF-8054-X; 11-Oxo-18,22-cycloergosta-6,8(14)-diene-3beta,5beta,9beta-23S-tetraol	<i>Emericella heterothallica</i> , <i>Aspergillus ustus</i>	R. Mizuno et al., Chem. Pharm. Bull. 43 (1995) 9-11; Sakai, K. et al., J. Antibiot. 47 (1994) 591-594; T. Hosoe et al., Chem. Pharm. Bull. 46 (1998) 850-2
			3a,5beta,8beta,11beta-Diepidioxy-ergost-22E-en-12-one	<i>Trametes orientalis</i>	Chen, H.L. et al. J. Chin. Chem. Soc. (Taipei) 42 (1995) 97-100
<b>C<sub>21</sub>H<sub>33</sub>O<sub>3</sub></b>	333.24171	4	Tintinnadiol	<i>Mycena tintinnabulum</i> (Basidiomycete)	Engler, M. et al., Phytochemistry. 49 (1998) 2591-3
			Demethylcisterol A3; O-Demethyl-17-methylcisterol	<i>Agrocybe chaxingu</i> (Basidiomycete)	Kawagishi, H. et al., Heterocycles. 69 (2006) 253-258
			Verticipyrone	<i>Verticillium sp.</i>	Ui, H. et al., J. Antibiot. 59 (2006) 785-90
			Neosarcodonin O	<i>Sarcodon scabrosus</i>	Kamo, T. et al., Biosci. Biotechnol. Biochem. 68 (2004) 1362-5

Travail expérimental – Étude du métabolome secondaire de *F. graminearum*

Formula	<sup>12</sup> C m/z	AB hits	Putative metabolite	Producing Organism	Reference
<b>C<sub>32</sub>H<sub>48</sub>O<sub>5</sub>Na</b>	535.33879	6	Ganoderic acid X	<i>Ganoderma lucidum</i>	Toth, J.O. et al., J. Chem. Research (1983) 2722-2787;
			3-O-Acetyl-16a-hydroxydehydrotrametenolic acid	<i>Poria cocos</i>	Tai, T. et al., Phytochem. 40 (1995) 225-231
			Dehydroeburicoic acid	<i>Antrodia cinnamomea</i>	Yang, S.W. et al., Phytochem. 41 (1996) 1389-1392
			Versisponic acid-B	<i>Laetiporus versisporus</i>	Yoshikawa, K. et al., Chem. Pharm. Bull. 48 (2000) 1418-21
			Colossolactone-B	<i>Ganoderma colossum</i> (Basidiomycete)	Kleinwaechter, P. et al., J. Nat. Prod. 64 (2001) 236-239
			Tyromyic acid C	<i>Tyromyces fissilis</i> (Basidiomycete)	Quang, D. N. et al., J. Nat. Prod. 67 (2004) 148-151
<b>C<sub>30</sub>H<sub>45</sub>O<sub>2</sub></b>	437.34157	1	Ganoderal	<i>Ganoderma concinna</i> , <i>G. pfeifferi</i>	Gonzalez, A.G. et al., J. Nat. Prod. 65 (2002) 417-421
			<b>C<sub>28</sub>H<sub>40</sub>O<sub>4</sub>Na</b>	463.28185	2
			2-Acetoxy-3-geranylgeranyl-1,4-dihydroxybenzene	<i>Suillus granulatus</i>	C. Tringali et al., J. Nat. Prod. 52 (1989) 941-7
ESI -	<b>C<sub>18</sub>H<sub>21</sub>O<sub>6</sub></b>	333.13502	8'-Epihydroxy-zearalenone; F-5-4	<i>Gibberella zeae</i> ( <i>fusarium graminearum</i> )	Cole/Cox, 907
				<i>Gibberella zeae</i> ( <i>fusarium graminearum</i> )	Cole/Cox, 907; Pohland, A.E. et al., Pure Appl. Chem., 54 (1982) 2219;
			8'-Hydroxy-zearalenone; F-5-3	<i>Fusarium roseum</i>	Turner II, 176
			3'-Hydroxyzearalenone	fungus IFM 52672	Hosoe, T. et al., J. Antibiot. 57 (2004) 573-8; Hosoe, T. et al., Heterocycles 63 (2004) 2581-2589
			Dihydroepihevadride	<i>Aigialus parvus</i> ( <i>marine fungus</i> )	Isaka, M. et al., J. Org. Chem. 67 (2002) 1561-6
			Aigialomycin D	<i>Piloderma croceum</i> ( <i>Corticium croceum</i> )	Gill/Steglich, 181 (fo 400); Rives-Arnau, V., Macromolecules, 18 (1985) 2088;
			2,4,6,8,10,12-Tetradecahexaene-1,14-dioic acid; Corticrocin	Not specified	Thomson II, 165; Thomson III, 89
			Tetrahydrocypaquinone	Not specified	Thomson III, 264
			Hibiscoquinone B	<i>Phaeoacremonium sp.</i>	Turner II, 110; Reategui, R. F. et al., J. Nat. Prod. 69 (2006) 113-117
			5'-Formyl-2'-hydroxy-4'-methoxy-(E,E)-sorbophenone; 1-(5'-Formyl-2'-hydroxy-4'-methoxy)-2,4-hexadien-1-one	Not specified	Gill/Steglich, 121 (fo 278)
Torachryson	<i>Penicillium chrysogenum</i>	Miller, R.F. et al., J. Antibiot. 48 (1995) 520-1			
Sorrentanone	<i>Pyricularia oryzae</i>	Nukina, M. Rec. Res. Devel. Agric. Biol. Chem. 2 (1998) 411-417			
Dihydropyriculone	<i>Fusarium sp.</i>	Xang et al., 2011 (Bioorganic & Medical Chemistry Letters 21, 2011, p 1887-1889)			
<b>C<sub>18</sub>H<sub>21</sub>O<sub>5</sub></b>	317.14018	1	Fusariumin		



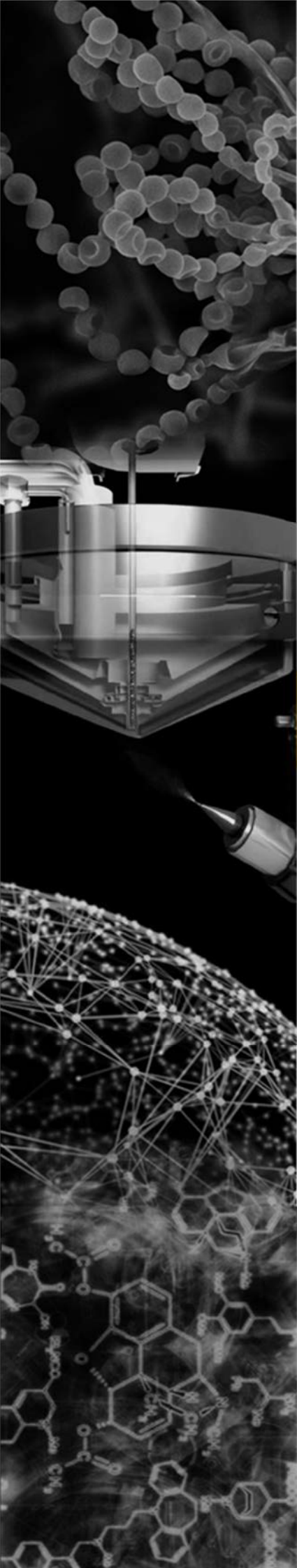
Travail expérimental – Étude du métabolome secondaire de *F. graminearum*

Formula	<sup>12</sup> C m/z	AB hits	Putative metabolite	Producing Organism	Reference
<b>C<sub>25</sub>H<sub>35</sub>O<sub>5</sub></b>	415.24949	5	Ophiobolin L	<i>Helminthosporium marydis</i>	E. Li et al., J. Nat. Prod. 58 (1995) 74
			S-15183a	<i>Zopfella inermis</i>	Keita, K. et al., J. Antibiot. 54 (2001) 415-20
			Blazeispirol G	<i>Agaricus blazei</i>	Hirovani, M. et al., Phytochemistry 61 (2002) 589-595
			17-Dehydroxyhalorosellinic acid	<i>Halorosellinia oceanica</i> (marine fungus)	Chinworrungsee, M. et al., J. Chem. Soc. Perkin 1 (2002) 2473-6
			Neomangicol C	<b>marine <i>Fusarium</i> sp.</b>	Renner, M.K. et al., J. Org. Chem. 63 (1998) 8346-54
<b>C<sub>18</sub>H<sub>31</sub>O<sub>4</sub></b>	311.22318	4	Filoboletic acid; (9Z,11E)-8,13-dihydroxyoctadeca-9,11-dienoic acid	<i>Filoboletus</i> sp.	Simon, B. et al., Phytochemistry 36 (1994) 815-816
			Piliferolide C	<i>Ophiostoma piliferum</i>	Ayer, W.A. et al., Heterocycles 39 (1994) 561-9
			Hygrophorone A12	<i>Hygrophorus persoonii</i> (Basidiomycete)	Lübken, T. et al., Phytochemistry 65 (2004) 1061-71; Lübken, T. et al., J. Mass Spectrom. 41 (2006) 361-371
			6-Hydroxy-4-oxo-11-octadecenoic acid	<i>Hygrophorus discoxanthus</i> (Basidiomycete)	Gilardoni, G. et al., Nat. Prod. Comm. 1 (2006) 1079-1084
<b>C<sub>18</sub>H<sub>29</sub>O<sub>3</sub></b>	293.21255	6	(S,S)-Conocandin; A-32287	<i>Hormococcus conorum</i>	Banfi, L. et al., J. Org. Chem. 1987, 52, 5452-7; Helv., 59, 2506, 1976
			11-Keto-9(E),12(E)-octadecadienoic acid	<i>Trichoderma</i> sp.	Shinohara, C. et al., J. Antibiot. 52 (1999) 171-4; Chikanishi, T. et al., J. Antibiot. 52 (1999) 797-802
			9-Oxo-10,12-octadecadienoic acid	<i>Clitocybe clavipes</i> (Basidiomycete)	Kawagishi, H. et al., J. Nat. Prod. 65 (2002) 1712-1714
			Hygrophorone G12	<i>Hygrophorus species</i> (Basidiomycete)	Lübken, T. et al., Phytochemistry 65 (2004) 1061-71; Lübken, T. et al., J. Mass Spectrom. 41 (2006) 361-371
			(1S,2R)-3-Oxo-2-(2Z-pentenyl)cyclopentane-1-octanoic acid	<b><i>Fusarium oxysporum</i></b>	Miersch, O. et al., Phytochem. 50 (1999) 517-23
			(1S,2S)-3-Oxo-2(2Z-pentenyl)cyclopentane-1-octanoic acid	<b><i>Fusarium oxysporum</i></b>	Miersch, O. et al., Phytochem. 50 (1999) 517-23
<b>C<sub>25</sub>H<sub>35</sub>O<sub>3</sub></b>	383.25961	12	6-Epiophiobolin I	<i>Drechslera oryzae</i>	Sugawara, F. et al., J. Org. Chem. 1988, 53, 2170-2 (Spec)
			8-Deoxyphiobolin J	<i>Drechslera oryzae</i>	Sugawara, F. et al., J. Org. Chem. 1988, 53, 2170-2 (Spec)
			Ophiobolin I; Cochliobolin-C	<i>Drechslera oryzae</i> ; <i>Cochliobolus miyabeanus</i>	Sugawara, F. et al., J. Org. Chem. 1988, 53, 2170-2 (Spec); Li, E. et al., J. Nat. Prod. 58 (1995) 74
			6-Epiophiobolin-K	<i>Aspergillus ustus</i>	Singh, S. B. et al., Tetrahedron 47, 6931-8 1991
			Ophiobolin-K	<i>Aspergillus ustus</i>	Singh, S. B. et al., Tetrahedron 47, 6931-8 1991
			Ophiobolin M	<i>Cochliobolus heterostrophus</i>	Anke, H. et al., Current Org. Chem. 1 (1997) 361-374
			6-Epiophiobolin M	<i>Cochliobolus heterostrophus</i>	Anke, H. et al., Current Org. Chem. 1 (1997) 361-374
			6-Epi-3-anhydrophiobolin B	<i>Cochliobolus heterostrophus</i>	Shen, X.Y. et al., J. Nat. Prod. 62 (1999) 895-897

Formula	<sup>12</sup> C <i>m/z</i>	AB hits	Putative metabolite	Producing Organism	Reference
			Variecolactone; EA-2	<i>Emericella purpurea</i> , <i>Emericella aurantio-brunnea</i>	Takahashi, H. et al., J. Nat. Prod. 62 (1999) 1712-3; Fujimoto, H. et al., Chem. Pharm. Bull. 48 (2000) 1436-41
			4,4,14a-Trimethyl-3-oxo-5a-pregna-7,9(11)-diene-20(S)-carboxylic acid	<i>Cladosporium sp.</i> , <i>Verticillium lecanii</i>	Grove, F.J. at al., J. Chem. Soc. I 1984, 1219-21
			Variecolactone; EA-2	<i>Emericella purpurea</i> , <i>Emericella aurantio-brunnea</i>	Takahashi, H. et al., J. Nat. Prod. 62 (1999) 1712-3; Fujimoto, H. et al., Chem. Pharm. Bull. 48 (2000) 1436-41
			4,4,14a-Trimethyl-3-oxo-5a-pregna-7,9(11)-diene-20(S)-carboxylic acid	<i>Cladosporium sp.</i> , <i>Verticillium lecanii</i>	Grove, F.J. at al., J. Chem. Soc. I 1984, 1219-21

**Table 1. Possible annotations for the new metabolites produced by *F. graminearum* based on AntiBase hits.** Chemical formulas obtained after the HPLS-HRMS analysis of *F. graminearum* cultures extracts were searched in AntiBase 2012. The results shown below are restricted to fungal metabolites. In bold are depicted the organisms that are phylogenetically closest to *F. graminearum*.





# Discussion

« The important thing  
is to never stop questioning. »

*Albert Einstein*



Les travaux réalisés au cours de cette thèse ont permis de mieux appréhender les métabolomes secondaires de trois champignons toxigènes. Les résultats obtenus ont été discutés dans les articles précédents. La section qui suit a pour objectif d'approfondir ces idées et de discuter plus précisément des perspectives qui permettraient de compléter, à court ou long terme, ce travail de thèse.

Afin de considérer au mieux les résultats de cette étude, une première partie de la discussion est consacrée à l'analyse critique de la méthode employée. Cette section porte tout d'abord sur les outils employés dans le cadre de cette thèse, c'est-à-dire sur la stratégie du marquage isotopique ainsi que sur les outils informatiques. Puis, une discussion à plus large échelle traite de l'étude métabolomique des métabolomes secondaires fongiques. Enfin, puisqu'une part importante du projet a consisté en l'élucidation structurale de peptides, une section est dédiée à la discussion quant aux mécanismes de fragmentation peptidique.

Dans un deuxième temps, les résultats obtenus au cours de ce projet sont discutés. Certains résultats et propos abordés dans les articles seront rappelés afin d'assurer la cohérence du discours. Une première partie évoque les différentes perspectives qui permettraient d'investiguer la part importante de métabolites secondaires fongiques encore inconnus à ce jour. Puis, une section est consacrée à l'étude de *F. graminearum* et à la découverte de deux nouvelles fusaristatines. Dans un troisième temps, les résultats quant aux métabolomes secondaires des espèces de *Penicillium* sont discutés. Plus particulièrement, des hypothèses et perspectives quant aux nouvelles fungisporines sont détaillées.

Au cours de cette thèse, la comparaison des métabolomes secondaires d'espèces fongiques a mis en évidence l'intérêt de la classification des moisissures selon des critères métaboliques. Néanmoins, la détection d'un métabolite secondaire inattendu chez *P. verrucosum* a témoigné de certaines limites inhérentes à une telle classification. Dans une troisième partie de la discussion, un système alternatif de classification métabolique des champignons filamenteux est imaginé.



## 1. ANALYSE CRITIQUE DE LA MÉTHODE

---

### 1.1. Marquage isotopique du substrat fongique

#### 1.1.1. Une méthode adaptable à de multiples sujets d'étude

Dans cette étude, il a été démontré que le couplage du marquage isotopique du substrat fongique et de l'analyse par spectrométrie de masse haute résolution permet de détecter spécifiquement les métabolites secondaires sécrétés par un champignon. Grâce au double marquage isotopique, ces composés peuvent être caractérisés chimiquement et leur élucidation structurale est ainsi facilitée. Appliquée à d'autres moisissures, couplée à diverses méthodes de culture, d'extraction et de séparation, cette stratégie analytique peut être utilisée pour l'étude des métabolomes secondaires de tous les champignons filamenteux dont le substrat peut être marqué. Pour répondre à des problématiques économiques et environnementales multiples, diverses espèces de plantes peuvent être cultivées et marquées dans des enceintes hermétiques telles que celles mises au point par le CEA de Cadarache. Les champignons impliqués dans leur destruction, qu'ils soient pathogènes ou saprophytes, peuvent ensuite être cultivés sur leurs substrats naturels.

Au-delà des études de toxicologie, cette méthode peut être utilisée pour la recherche de composés d'intérêt industriels ou pharmaceutique. De plus, en supprimant l'étape d'infestation de la plante par un microorganisme, cette méthode peut être dérivée pour l'étude des métabolomes secondaires végétaux dans le cadre d'études de toxicité ou de recherches de produits à vertus thérapeutiques.

#### 1.1.2. Limites de la méthode

Bien qu'elle permette de surpasser certaines limites inhérentes à l'étude de substances naturelles, la stratégie de double marquage isotopique du substrat fongique présente un biais. En effet, la plante et le champignon qui la colonise étant tous deux entièrement marqués, il est difficile de discriminer les métabolites végétaux des métabolites fongiques. Certes, la comparaison des échantillons infestés et d'un témoin constitué uniquement de la plante permet de déterminer quelles sont les molécules liées directement à l'infestation de la plante par le champignon. Mais les molécules marquées pourraient aussi bien être produites par le champignon que par la plante en réponse à l'infection fongique. Dans le cadre de problématiques environnementales et toxicologiques, l'origine de nouveaux composés toxiques n'a pas d'importance tant qu'il est démontré que leur synthèse est stimulée par un contact entre l'hôte et le pathogène. Ce biais ne nuit donc pas à la significativité des résultats présentés ici.

Dans le cas où l'origine d'un métabolite donné est recherchée, plusieurs analyses sont envisageables. Par exemple, des analyses d'imagerie par spectrométrie de masse peuvent permettre de cartographier les concentrations relatives du métabolite de part et d'autre de la zone de contact entre le champignon et la plante. La comparaison des gradients de concentration entre le début du contact et un temps ultérieur indique le sens de diffusion de la molécule et le lieu de production du métabolite secondaire peut être déterminé. Une autre solution consiste en la détermination de l'origine génétique du métabolite étudié. Une fois le gène identifié chez la plante ou chez le champignon, la stimulation de sa transcription en l'absence de contact hôte-pathogène permet de



valider l'organisme producteur. De la même manière, l'appartenance d'un nouveau métabolite à une famille de métabolite secondaire connue permet d'émettre des hypothèses quant à son origine.

L'étude des métabolomes secondaires fongiques nécessite une prise en compte rigoureuse de la nature du substrat puisque les moisissures adaptent la synthèse des métabolites secondaires en fonction de leur environnement. Malheureusement, le double marquage isotopique présenté dans cette étude est applicable uniquement aux substrats cultivables dans un temps relativement court et dans des enceintes hermétiques de taille relativement réduite. Il n'est donc pas envisageable de générer des arbres ou des animaux marqués. Dans le cadre de cette thèse, l'étude de *P. nordicum* a donc été réalisée sur un substrat différent de son substrat naturel. Afin de prendre en compte l'influence de celui-ci, les métabolites secondaires observés lors de sa croissance sur blé ont été recherchés dans des cultures de *P. nordicum* sur viande. Ainsi, il est possible de caractériser les métabolites secondaires produits à la fois sur blé et sur viande. Néanmoins, il n'est pas possible de bénéficier des avantages du double marquage isotopique pour la caractérisation des métabolites produits spécifiquement lors de l'infestation de viande par *P. nordicum*.

### **1.2. Les outils bioinformatiques**

#### **1.2.1. CFM-ID**

Le logiciel en ligne CFM-ID est un outil bioinformatique qui permet de simuler des spectres MS/MS à partir de structures de molécules<sup>1</sup>. Bien qu'il ne constitue en aucun cas une preuve d'identification, ce logiciel a été utilisé dans cette thèse afin de soutenir les hypothèses structurales quant aux métabolites secondaires connus détectés mais dont ni les standards, ni les spectres MS/MS référencés n'étaient disponibles.

CFM-ID permet également la confrontation d'un spectre MS/MS inconnu avec les spectres MS/MS théoriques de toutes les molécules d'une base de donnée. Des tests réalisés dans le cadre du concours du CASMI (Critical Assessment of Small Molecule Identification) ont montré que pour ce type de recherche, CFM-ID identifie correctement une molécule dans seulement 15% des cas<sup>2</sup>. En revanche, en couplant les prédictions de CFM-ID avec celles de MetFrag, un deuxième outil de simulation de spectres MS/MS, et en utilisant les métadonnées des études métabolomiques pour filtrer les propositions incohérentes (le contexte biologique par exemple), il a été possible d'identifier correctement 93% des spectres testés<sup>2</sup>. Ces outils peuvent donc être une d'une aide précieuse lors des étapes de déréplication.

Dans le cadre de cette étude, certaines limites de l'utilisation de CFM-ID ont été rencontrées. Il a notamment été remarqué que la qualité de la simulation des spectres pouvait varier en fonction du temps. Alors que le spectre théorique de l'ochratoxine B était tout à fait cohérent avec le spectre MS/MS expérimental du composé standard à un temps donné, il pouvait être modifié considérablement lors d'une nouvelle analyse quelques mois plus tard. Ceci peut être expliqué par l'évolution de l'apprentissage automatique (Machine Learning) sur lequel est fondé ce logiciel. En effet, lors de son développement, CFM-ID est alimenté avec les structures de composés de référence et leurs spectres MS/MS générés à trois niveaux d'énergie différents. Lorsqu'un utilisateur souhaite obtenir le spectre théorique d'une molécule inconnue de CFM-ID, le logiciel génère les spectres MS/MS *in silico* selon les liens systématiques qu'il a faits lors de son apprentissage entre motifs de fragmentation et

sous-structures. L'apprentissage automatique a néanmoins l'avantage de considérer des mécanismes de fragmentation n'ayant pas été soulignés ni expliqués à l'heure actuelle.

Une autre limite de l'utilisation de CFM-ID en ligne réside dans le fait que le logiciel a réalisé son apprentissage automatique à partir de spectres MS/MS générés sur un spectromètre de masse de type Q-TOF. Or, chaque type de technologie fournit des spectres MS/MS spécifiques et il est important de comparer entre eux des spectres MS/MS provenant d'un même type d'appareils. Les utilisateurs de CFM-ID désirant comparer les spectres théoriques avec des spectres expérimentaux générés sur d'autres types de spectromètres de masse doivent donc considérer avec précaution les résultats du logiciel. Le moyen le plus adéquat d'utiliser les outils de simulation de spectres MS/MS fondés sur un apprentissage automatique est d'enrichir le logiciel en interne avec des spectres MS/MS standards générés sur les mêmes spectromètres de masse que ceux qui sont utilisés dans les analyses métabolomiques. Néanmoins, cette démarche est contraignante puisqu'elle nécessite l'annotation manuelle de tous les spectres MS/MS fournis au logiciel.

Enfin, les outils de simulation de spectres MS/MS n'interviennent que dans l'identification de composés connus présents dans les bases de données. L'enjeu majeur de l'analyse des métabolomes secondaires fongiques réside dans la caractérisation de molécules inconnues.

### 1.2.2. Les réseaux moléculaires

Le système du GNPS Molecular Networking a été développé notamment dans le but d'assister les analyses métabolomiques lors de la caractérisation de molécules inconnues. Grâce à cet outil, de nouveaux métabolites secondaires ont été identifiés lors de cette thèse. Les avantages du logiciel du GNPS ont été détaillés précédemment, c'est pourquoi ils ne seront pas de nouveau mentionnés dans cette partie de la discussion consacrée aux difficultés rencontrées dans le cadre de l'utilisation de ce logiciel.

Lors de la génération de réseaux moléculaires, et en particulier lors de l'optimisation des paramètres, certaines difficultés ont été rencontrées en raison du manque d'adaptabilité des paramètres à chacun des spectres. Notamment, la précision de masse doit être spécifiée en dalton. Or, cette unité ne permet pas de considérer la variation de l'écart de masse en fonction du rapport  $m/z$  du composé. Afin de respecter les données analytiques, il serait préférable que la précision de masse soit renseignée en partie par million (ppm). De plus, les extraits naturels sont composés de molécules de structures très variables. Le paramètre concernant le nombre d'ions fragments communs à considérer pour établir un lien entre deux spectres MS/MS est donc délicat à établir (voir Travail expérimental, Figure 34). L'une des solutions qui permettrait de s'affranchir de ce problème est de considérer ce paramètre non pas en valeur fixe mais en pourcentage d'ions fragments totaux. Ainsi, le logiciel évaluerait la fragilité des molécules en considérant le nombre d'ions fragments présents dans leurs spectres MS/MS. Puis, il déterminerait le nombre d'ions fragments minimum nécessaires pour lier deux molécules en fonction d'un pourcentage de ce nombre total d'ions fragments.

Une deuxième limite spécifique de l'analyse d'extraits naturels réside dans le fait que les concentrations des différents composés sont extrêmement variables. Comme la sensibilité d'un spectromètre de masse est stable, la qualité des spectres MS/MS peut être médiocre pour certains métabolites. Le paramètre concernant l'abondance minimum des ions fragments pour les différencier du bruit de fond est donc également délicat à établir (voir Travail expérimental, Figure 34). Afin de ne pas prendre en considération les signaux parasites dans la comparaison de deux spectres MS/MS, une

étape d'extraction des ions fragments au sein de chaque spectre MS/MS pourrait être mise en œuvre. La discrimination du bruit de fond pourrait être effectuée en comparant les spectres MS/MS issus de chaque scan sous un même pic chromatographique. Les ions dont le rapport  $m/z$  est constant (à une précision de masse donnée) seraient alors considérés comme des ions fragments et les ions d'apparition aléatoire seraient rejetés. Cette étape pourrait intervenir avant l'étape de regroupement des spectres MS/MS moyens évoquée ci-après.

Avant de comparer les données entre elles, le GNPS regroupe en un spectre MS/MS moyen tous les spectres MS/MS similaires issus de précurseurs de mêmes rapports  $m/z$ . Cette action est fondée sur une approche d'archives spectrale qui ne prend pas en considération les temps de rétention relatifs aux spectres MS/MS, ni les fichiers distincts dont ont été extraits les spectres (voir Travail expérimental, Partie 2, Section 2.2). Afin de générer des réseaux représentatifs de l'ensemble des molécules analysées et donc de permettre la discrimination des différents isomères, il serait judicieux d'intégrer un processus d'extraction de pics préalablement au regroupement des spectres MS/MS similaires.

Le calcul de similarité entre deux spectres MS/MS est réalisé par le GNPS avec un score cosinus donnant autant de poids à l'intensité des ions fragments qu'à leur rapport  $m/z$  (voir Travail expérimental, Partie 2, Section 2.2). Or deux molécules très similaires ne différant que par la localisation ou l'ajout d'un groupement fonctionnel peuvent présenter des ions fragments de même rapport  $m/z$  mais d'intensité très variables. Afin de surpasser cette limite, la similarité entre deux spectres MS/MS devrait être évaluée par la seule considération des masses des ions fragments et des pertes de neutres entre eux. En revanche, l'intensité des ions fragments devrait être considérée pour la détection de spectres MS/MS identiques provenant d'une même molécule. Le score cosinus du GNPS pourrait alors être utile pour la génération des spectres moyens, pour repérer des molécules identiques provenant de différentes analyses (deux échantillons différents, ou un échantillon et un standard), ou pour interroger les bases de données de spectres MS/MS. Enfin, le degré de similarité entre deux molécules est établi par le GNPS sans considération de la spécificité de leurs structures communes. Deux molécules présentant une perte d'eau en commun (représentative de la présence d'un groupement hydroxyle par exemple) seront jugées aussi similaires que deux molécules présentant un fragment à  $m/z$  120,0804 en commun (pouvant représenter la présence d'une phénylalanine). Il pourrait donc être pertinent d'inclure dans le calcul de similarité entre deux spectres MS/MS un score donnant plus de poids aux ions fragments de faible rapport  $m/z$  (en fonction d'un pourcentage par rapport à la masse du parent par exemple) ou aux petites pertes de neutres. Il serait appréciable que ce paramètre soit optionnel puisque l'observation de petites pertes de neutre peut également être informative selon les études.

### 1.2.3. Développement d'un nouvel outil bioinformatique

Suite à ces observations, une collaboration a été établie avec le CEA de Saclay afin de développer un outil informatique qui puisse s'affranchir de ces limites. Cette collaboration a pour objectif de lier les expertises en spectrométrie de masse ainsi que l'expérience des réseaux métaboliques et les connaissances en terme de développement d'outils bioinformatiques.

Alors que les réseaux moléculaires du GNPS fonctionnent via une vectorisation des spectres MS/MS, l'outil en développement dans le cadre de la collaboration avec le CEA est fondé sur la

comparaison de motifs. Dans un premier temps, l'extraction des données consiste en une étape d'extraction des pics MS selon leurs profils chromatographiques. Cela permet à la fois la discrimination du bruit de fond présent dans le signal MS et la discrimination des différents isomères. Puis, les spectres MS/MS présents sous ces pics sont extraits et moyennés (Figure 36A). Dans un deuxième temps, pour chacun des spectres MS/MS moyens, les ions fragments sont extraits et un graphe orienté est généré. Dans celui-ci, les sommets représentent les différents ions fragments et les arrêtes sont annotées avec les différences de rapport  $m/z$  qui sont calculées entre tous les ions du spectre MS/MS deux à deux (Figure 36B). Les graphes correspondant à chacun des spectres MS/MS sont ensuite comparés entre eux et des motifs communs sont mis en évidence (Figure 36C). Enfin, les motifs sont annotés chimiquement (Figure 36D). Le logiciel est en cours de développement et des algorithmes sont élaborés afin de surpasser les différentes limites rencontrées lors de la création de réseaux moléculaires. Bien que l'outil actuel ne soit fondé que sur la considération des pertes de neutres, il est envisagé d'intégrer la considération des valeurs  $m/z$  des ions fragments.

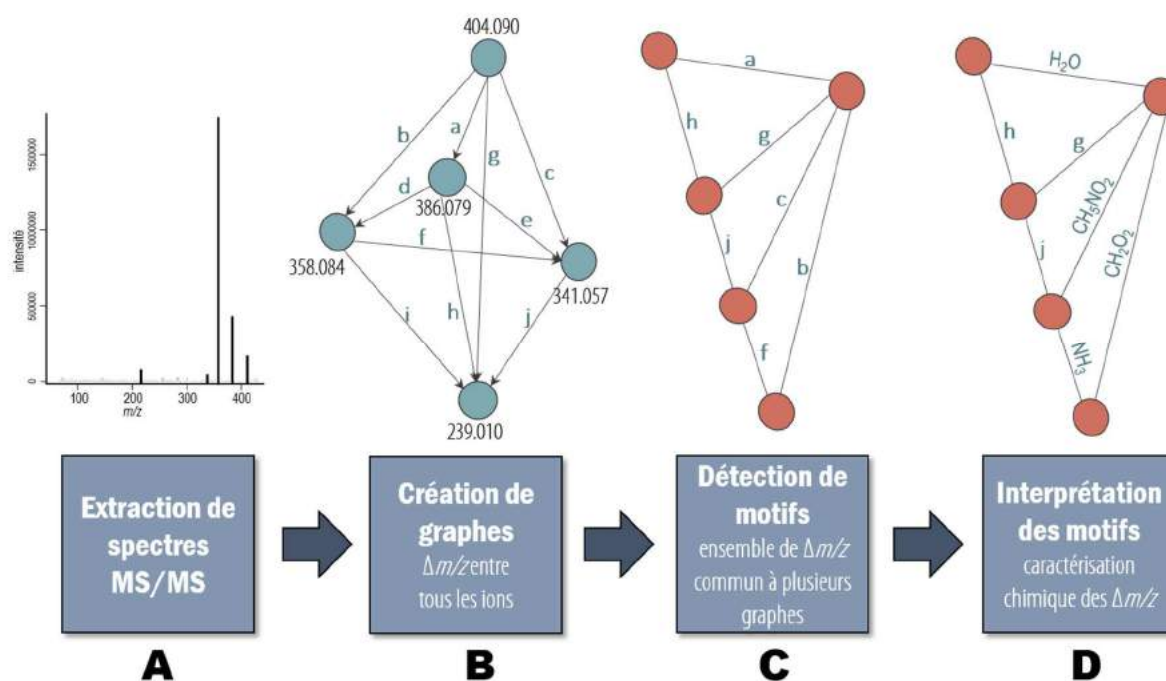


Figure 36. Protocole d'extraction et de traitement des données de l'outil en développement.

### 1.3. Analyse critique de l'étude métabolomique des champignons filamenteux

A travers cette étude, certaines limites liées aux analyses métabolomiques de manière générale ont été rencontrées. Bien que le terme « analyse globale » soit parfois employé pour mentionner une méthode d'analyse non-ciblée, les techniques d'extraction, d'ionisation et d'analyse appliquées dans ce projet ne permettent pas la détection de la totalité des métabolites secondaires produits par le champignon. Certaines molécules ne sont pas extraites par l'acétate d'éthyle, les molécules volatiles sont préférentiellement séparées par chromatographie en phase gazeuse et certains composés ne sont pas ionisés en mode electrospray. Coupler différentes analyses réalisées avec des méthodes de préparations d'échantillon différentes et analysées avec différentes techniques permettrait de détecter davantage de métabolites secondaires.

Les champignons sont des microorganismes capables d'adapter leur métabolome secondaire en fonction de leur environnement afin de répondre à des besoins particuliers en terme de croissance ou de nutrition dans un milieu plus ou moins hostile. Il est donc préférable d'analyser les métabolomes secondaires fongiques lors de la culture du champignon sur un milieu proche de son substrat naturel. Malgré les efforts réalisés dans ce sens, il est rare que les conditions de croissances soient exactement identiques à celles rencontrées naturellement. Dans le cadre de cette thèse, le champignon phytopathogène *P. verrucosum*, responsable de la destruction des grains de blé lors de leur stockage, a été étudié sur grains de blé morts. Néanmoins, ces conditions ne reflètent pas les réelles conditions de croissance lors desquelles le champignon peut être confronté à d'autres microorganismes, à des variations de température, d'humidité ou de luminosité par exemple. Exposer des grains de blé marqués dans les conditions réelles de stockage ne permettrait pas de garantir que l'unique source de carbone et d'azote du champignon soit marquée en raison de potentielles contaminations du stock. De plus, les molécules analysées ne sauraient être spécifiquement attribuées à *P. verrucosum* et non à d'autres microorganismes présents dans le milieu. Pour toutes ces raisons, la détermination des métabolites secondaires produits par un champignon dans les conditions réelles de croissance est une tâche complexe.

Lors de l'analyse des produits naturels, il a été démontré que de nombreux métabolites naturels sont redécouverts, c'est-à-dire qu'après purification d'une molécule non annotée, les analyses structurales révèlent une structure d'ores et déjà caractérisée. Ces événements entraînent des pertes de temps importants et limitent la découverte de composés réellement inconnus. L'étape de déréplication est donc cruciale afin de cibler spécifiquement l'analyse des composés inconnus. Néanmoins, la faible disponibilité des composés standards ou de leurs spectres MS/MS limite cette étape. Dans la littérature, ce sont généralement les spectres RMN qui sont partagés lors de la découverte d'un nouveau composé, et seul le rapport  $m/z$  de l'ion précurseur est mentionné. Or, ce sont les spectres MS/MS ou les spectres UV qui sont nécessaires à l'identification rapide de composés présents en faible concentration dans un mélange complexe. Certaines technologies telles que la micro-RMN pourraient permettre de surpasser cette limite par l'analyse directe des molécules en sortie de colonne chromatographique, mais cet outil n'est pas encore très répandu. De plus, la sensibilité de ce type de technique peut poser question. Dans le cadre de cette thèse, les données MS/MS concernant tous les composés détectés ont donc été partagées. Les métabolites secondaires identifiés au niveau 3, c'est-à-dire ceux dont le spectre MS/MS est compatible avec la structure hypothétique mais dont nous n'avons aucune preuve (composés standards ou spectres référencés), sont à considérer avec précaution, en particulier concernant les composés n'ayant jamais été observés auparavant chez ces espèces.

Les technologies d'analyse de haute sensibilité et de haute résolution utilisées dans le cadre de projets métabolomiques permettent de générer un grand nombre de données à partir d'un seul échantillon. Aujourd'hui, les efforts de recherche sont orientés vers le développement d'outils bioinformatiques capables de traiter ces données non-interprétables par l'œil humain dans un temps raisonnable. Bien que ces outils récemment mis au point soient d'ores et déjà d'une aide précieuse, la plupart d'entre eux présentent des biais et demandent à être optimisés. Les résultats que fournissent ces outils devraient donc être considérés avec précaution. Les enjeux actuels sont de deux types. Premièrement, il est crucial de poursuivre le développement d'outils bioinformatiques et d'optimiser leurs analyses afin que leur fiabilité soit renforcée. Deuxièmement, il est important de sensibiliser les

utilisateurs à ces précautions, et de stimuler le développement d'expertise dans l'analyse des données brutes et de l'esprit critique vis-à-vis de résultats issus de ces outils bioinformatiques. Ce dernier point est notamment encouragé par les journaux de protéomique qui demandent des spectres annotés et des données brutes lors de la publication de séquences de nouvelles protéines élucidées par des outils de séquençage automatique.

#### 1.4. Discussion autour de la fragmentation peptidique

Lors de l'analyse structurale des nouveaux peptides, l'annotation des spectres de fragmentation a été réalisée sans l'aide d'outil de séquençage automatique. Ceci a permis, non seulement de s'affranchir des biais que peuvent présenter certains de ces outils, mais également de considérer tous les ions fragments sans se restreindre aux ions issus des fragmentations dites classiques (ions **a**, **b** et **y**). La connaissance et l'application des règles de fragmentation, qu'elles soient particulières aux peptides ou générales, sont en effet présentées comme des atouts pour une interprétation rigoureuse de spectres MS/MS et l'élucidation de nouvelles structures<sup>3</sup>.

L'interprétation manuelle des spectres MS/MS permet également de considérer avec précaution certains mécanismes décrits dans la littérature et d'opter pour l'un ou l'autre des multiples mécanismes parfois décrits pour une même fragmentation. Dans le cas des peptides par exemple, les spectres MS/MS d'adduits au sodium ont été interprétés suivant l'hypothèse que la présence de l'atome de sodium entraînait la formation d'un zwitterion et que la fragmentation était orientée par la charge. Des mécanismes de fragmentation peptidique classiques ont alors été appliqués afin d'annoter les différents ions fragments. Cette hypothèse a permis d'interpréter les différents spectres MS/MS. Dans le cadre de cette thèse, l'atome de sodium a semblé être lié à un acide aminé donné, puisqu'il était présent dans la composition de chacun des ions fragments ou de chacune des pertes de neutre contenant cet acide aminé. Dans certains cas, l'acide aminé concerné était situé en milieu de séquence. Dans d'autres cas, le sodium a semblé lié au niveau de l'acide aminé N terminal, quelle que soit sa nature. Les principes de fragmentation des adduits sodium décrits par Renner *et al.*<sup>4</sup> ou par Lin *et al.*<sup>5</sup> n'ont pas contribué à l'élucidation de la structure des tétrapeptides produits par *P. nordicum*. Si le sodium était capté par l'ensemble des hétéroatomes du peptide, il est probable que chaque ion dipôle serait observé sous deux formes différentes : par exemple, l'ion dipôle **b<sub>2</sub>-y<sub>2</sub>** dans lequel l'atome de sodium serait capté par l'ion **b<sub>2</sub>** et l'ion dipôle équivalent dans lequel l'atome de sodium serait capté par l'ion **y<sub>2</sub>**. Or, ce phénomène n'a pas été observé.

L'analyse des fungisporines hydrolysées connues et l'étude des spectres MS/MS des composés standards ont permis de mettre en évidence des ions fragments n'ayant à notre connaissance jamais été décrit précédemment. Des hypothèses quant aux mécanismes à l'origine de ces ions ont été émises sur la base des règles de fragmentation générales. Ces hypothèses ont été en partie appuyées par la réalisation de spectres MS<sup>3</sup> et MS<sup>4</sup> et la comparaison de fragmentations compétitives et consécutives (en modes CID et HCD respectivement). Ces nouveaux types d'ions fragments ont été observés chez tous les tétrapeptides analysés quelles que soient leurs compositions. Afin de justifier les mécanismes de fragmentation proposés et les structures proposées, plusieurs analyses pourraient être effectuées. Notamment, des expériences d'échange proton-deutérium pourraient permettre de vérifier le type d'acides aminés inclus dans les ions fragments. Le tryptophane et la tyrosine incluant un proton échangeable, un ion fragment possédant un seul de ces acides aminés verra son *m/z* décalé d'une unité dans le spectre de fragmentation du peptide dont tous les protons échangeables ont été

échangés. En revanche, un ion fragment composé d'une phénylalanine et d'une valine (sans proton échangeable) ne verra pas son rapport  $m/z$  modifié. L'échange de protons permet également de vérifier le degré de substitution d'un azote. Alors qu'une amine quaternaire (chargée positivement) ne présentera aucun proton échangeable, une amine protonée verra son  $m/z$  décalé (Figure 37).

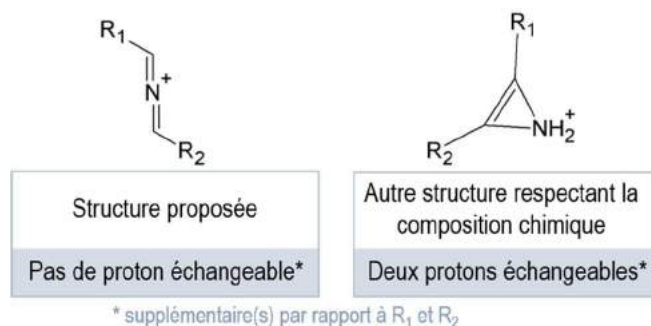


Figure 37. Utilisation de l'échange hydrogène/deutérium pour la validation d'hypothèses structurales.

Afin de justifier les mécanismes de fragmentation proposés, des calculs théoriques quantiques permettant d'évaluer les états de transitions pourraient également être réalisés. Dans le cas où les hypothèses formulées à propos de la fragmentation de ces ions seraient vérifiées, ces nouveaux types d'ions satellites pourraient participer à l'élucidation structurale de peptides et être intégrés dans les outils informatiques de séquençage peptidique. Des études supplémentaires permettraient de rechercher la présence d'ions équivalents lors de la fragmentation de chaînes peptidiques constituées de plus de quatre acides aminés.

## 2. ANALYSE DES RÉSULTATS DE L'ÉTUDE

### 2.1. Environ 80% de métabolites secondaires inconnus

Grâce au couplage du marquage isotopique du substrat fongique et de la spectrométrie de masse, 80, 70 et 29 métabolites secondaires totalement inconnus ont été détectés chez *P. verrucosum* et *P. nordicum* et *F. graminearum* respectivement. Ces chiffres représentent respectivement 81%, 76% et 78% de métabolites inconnus au sein des métabolomes de ces trois espèces toxigènes. Des études structurales sont à présent nécessaires pour caractériser ces nouvelles molécules. De plus, des études fonctionnelles permettraient de comprendre le rôle de chacun de ces métabolites secondaires. Deux approches sont notamment envisageables : l'une consacrée à l'étude de la toxicité de ces métabolites, et l'autre visant à étudier l'implication de ces composés dans le développement du champignon et donc dans l'infestation des denrées alimentaires.

Dans le cas des approches visant à étudier la toxicité de *P. verrucosum*, *P. nordicum* et *F. graminearum*, des expériences de mutation génétiques couplées à des purifications guidées par la toxicité pourraient permettre de repérer les métabolites impliqués dans une toxicité aiguë chez l'homme et l'animal. En effet, si un champignon dont les gènes responsables de la production des toxines connues ont été délétés présente des propriétés toxiques, cela souligne la présence d'autres composés toxiques qui n'avaient pas été observés auparavant. Une purification guidée par cet effet toxique permettrait ensuite de caractériser structurellement le métabolite responsable et d'étudier spécifiquement sa toxicité. Parfois, la toxicité peut être engendrée par un mélange de métabolites. Dans ce cas, il convient de repérer la plus petite fraction du métabolome à l'origine de cet effet, puis

de purifier séparément chacun des métabolites présents dans cette fraction. En prenant en compte leur concentrations relative, des tests réalisés à partir de différents mélanges pourraient permettre de mettre en évidence le mélange minimum nécessaire à la génération de cet effet. Cette approche est adaptée pour l'étude de la toxicité aiguë uniquement. Elle est en effet inenvisageable pour tester sur le long terme le grand nombre de mélanges qu'elle est susceptible d'exiger.

Pour l'étude de la toxicité chronique des nouveaux métabolites, des analyses *in silico* de modélisation moléculaire et de simulation des paramètres pharmacocinétiques pourraient tenter de prédire les étapes d'absorption, de distribution et d'interaction avec les différents éléments cellulaires pour envisager un potentiel effet toxique à long terme. Des études épidémiologiques peuvent également mettre en évidence la toxicité chronique de composés mais dans le cas d'un métabolome complexe comme celui des champignons, ce type d'étude ne permettrait peut-être pas de distinguer les métabolites responsables.

Pour étudier l'implication des nouveaux métabolites sur la colonisation des denrées alimentaires par *P. verrucosum*, *P. nordicum* et *F. graminearum*, des expériences de manipulation génétiques seraient adaptées. Notamment, des souches présentant chacune l'un des clusters de gènes inconnu délété pourraient être générées. Si l'une d'entre elles présente des difficultés développement, il pourra être suggéré que le cluster de gène délété soit impliqué dans la croissance du champignon. La comparaison des métabolomes secondaires de la souche sauvage et de la souche mutée pourrait ensuite mettre en évidence le métabolite précisément impliqué dans cette fonction. De plus, si l'un des mutants ne parvient pas à se développer sur grains de blé (dans le cas de *P. verrucosum* ou de *F. graminearum* par exemple), mais possède une croissance normale (équivalente à la souche sauvage) sur milieu de culture standard, il est possible que ce métabolite soit spécifiquement impliqué dans la colonisation du milieu végétal. Ces observations pourraient ensuite être confirmées par l'insertion du gène de ce métabolite chez un champignon connu pour son incapacité à se développer sur grain de blé afin d'observer un potentiel transfert de capacité de développement sur céréales.

La caractérisation des métabolites secondaires impliqués dans le pouvoir pathogène de *P. verrucosum*, *P. nordicum* et *F. graminearum* permettrait de contrôler les denrées alimentaires et de protéger les populations. D'autre part, la connaissance des métabolomes secondaires de ces champignons permettrait de mettre en place des mesures ciblées pour limiter leur production de toxines tout en diminuant l'utilisation de pesticides.

## **2.2. Le métabolome secondaire de *F. graminearum***

### **2.2.1. Des métabolomes secondaires différents selon la souche et l'environnement**

*F. graminearum* est un champignon qui infeste les céréales lors de leur croissance. Dans le cadre de cette thèse, les métabolomes secondaires de deux souches de *F. graminearum* ont été investigués. Lors de la croissance de la souche de référence PH1 (INRA 605) *in vitro* sur grains de blé secs, 37 métabolites secondaires ont été caractérisés. Seuls 16 de ces métabolites ont été retrouvés lors de l'infestation artificielle d'épis de blé en croissance (*in planta*) par cette souche. En revanche, 26 des 37 métabolites secondaires ont été produits *in planta* par la souche INRA DGEC, isolée de champs de blé naturellement infestés. La différence métabolique entre les deux souches lorsqu'elles sont cultivées *in planta* pourrait s'expliquer par la perte de capacité à produire des métabolites au fur et à mesure de la croissance *in vitro* de la souche de référence PH1. Le fait que l'ensemble des 37



métabolites secondaires n'ait pas été produit par les deux souches lors de leur culture *in planta* pourrait être expliqué soit par leur présence dans des concentrations trop faibles pour qu'ils soient détectés, soit par une cinétique d'apparition différente, soit par des interactions plante-champignon modifiant le métabolome secondaire et limitant la production de certains composés fongiques.

La caractérisation des métabolites secondaires a été réalisée à partir de cultures *in vitro*, mais il est probable que *F. graminearum* produise des métabolites secondaires supplémentaires *in planta* pour faire face au métabolisme de la plante ou pour permettre sa dissémination au sein d'un organisme complexe. Afin de détecter spécifiquement ces composés et de déterminer leurs formules chimiques, une étude visant à infester les pieds de blé marqués aux isotopes stables au cours de leur croissance dans les chambres hermétiques est envisagée. La comparaison des métabolomes secondaires de *F. graminearum* obtenus *in vitro* et *in planta* permettrait ainsi de mettre en évidence les métabolites spécifiquement impliqués dans l'interaction avec la plante (messagers, défenses, attaques de l'hôte). L'objectif de ce projet est également d'investiguer l'aspect cinétique de la production des métabolites secondaires par *F. graminearum* en étudiant les composés synthétisés à différents temps d'infestation.

Lors de l'analyse de blé naturellement infesté par *F. graminearum*, il a été remarqué la présence systématique de souches de *F. culmorum*. Ce champignon étant également impliqué dans les fusarioses, les métabolites secondaires caractérisés précédemment ont été recherchés chez ces différentes souches. Il a été observé que les métabolomes de *F. culmorum* étaient très similaires à celui de la souche DGEC de *F. graminearum*. Cette remarque est en accord avec la proximité phylogénique très forte entre les deux espèces. Afin d'étudier plus précisément le métabolome secondaire de *F. culmorum*, une analyse non ciblée sur substrat marqué serait aussi pertinente.

### 2.2.2. Découvertes de nouvelles fusaristatines chez *F. graminearum*

La génération de réseaux moléculaires a permis de mettre en évidence des similarités de spectres MS/MS entre la fusaristatine A et deux métabolites secondaires inconnus. Le spectre MS/MS de la fusaristatine A a été interprété et les mécanismes de fragmentation ont été transposés pour élucider la structure de ces nouveaux composés. Ainsi, la fusaristatine C a présenté une structure correspondant à celle de la fusaristatine A dont une alanine remplace la déshydroalanine. La fusaristatine D est quant à elle modifiée au niveau de la chaîne polycétide. À partir de la voie de biosynthèse de la fusaristatine A proposé par Sorensen *et al.*, il est possible d'émettre plusieurs hypothèses quant à l'origine biosynthétique de ces deux nouveaux métabolites secondaires<sup>6</sup>. Sorensen *et al.* supposent que l'incorporation de la déshydroalanine dans la fusaristatine A soit orchestrée par le premier module de la NRPS7 (Figure 38, encadrement vert). La déshydroalanine pourrait être générée à partir d'une alanine ou être issue de la déshydratation d'une sérine<sup>7</sup> par d'autres enzymes présentes dans le même cluster de gènes.

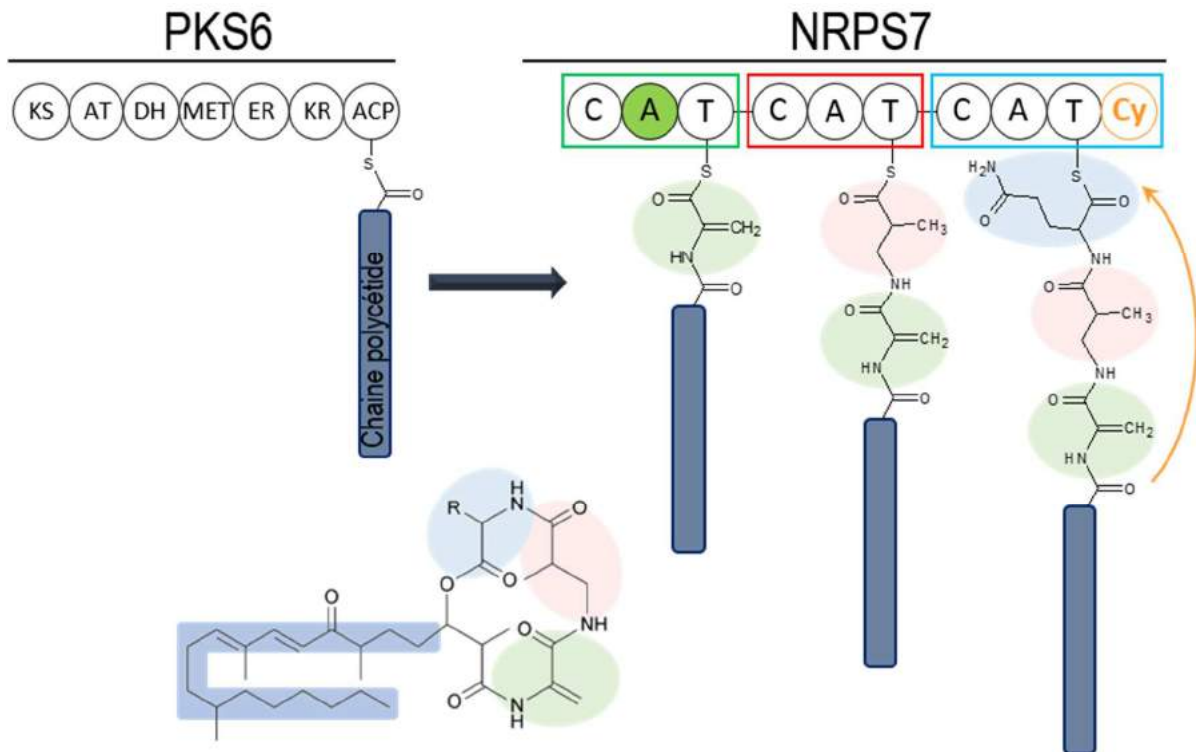


Figure 38. Voie de biosynthèse de la fusaristatine A proposée par Sorensen et al.<sup>6</sup> R = chaîne latérale de la glutamine

La présence d'une chaîne polycétide différente entre la fusaristatine A et la fusaristatine D suggère qu'une PKS différente de la PKS6 intervienne dans la synthèse de ce nouveau métabolite secondaire. Ainsi, la fusaristatine D pourrait être synthétisée à partir d'un cluster de gènes différent de celui à l'origine de la synthèse de la fusaristatine A.

Trois hypothèses sont suggérées quant à la voie de biosynthèse de la fusaristatine C :

(1) *Synthèse de la fusaristatine C via le même cluster de gènes que celui impliqué dans la synthèse de la fusaristatine A* : La première hypothèse consiste à suggérer que la fusaristatine C et la fusaristatine A soient issues du même cluster de gènes. Dans ce cas, le premier module de la NRPS7 est spécifiquement impliqué dans l'incorporation d'une alanine. L'enzyme responsable de la déshydrogénation de l'alanine lors de la synthèse de fusaristatine A intervient donc nécessairement après l'intégration de l'alanine dans le lipopeptide. Cette étape de déshydrogénation peut être évitée par des phénomènes d'itération, générant la fusaristatine C (Figure 39 Aa1).

(2) *Synthèse de la fusaristatine C via un cluster de gènes différent* : De multiples hypothèses impliquent que la fusaristatine C soit générée via un autre cluster de gènes que celui impliqué dans la synthèse de la fusaristatine A. Si l'on reprend l'hypothèse précédente, il est possible qu'un cluster de gènes soit très similaire à celui de la fusaristatine A mais se distingue par une absence ou une inactivation de la déshydrogénase (Figure 39 Aa2).

Dans le cas où la déshydroalanine de la fusaristatine A est issue de la déshydratation d'une sérine, deux situations sont probables. D'une part, la transformation de la sérine en déshydroalanine peut intervenir avant l'intégration de l'acide aminé par le premier module de la NRPS7. Dans ce cas, le premier domaine d'adénylation est spécifique de l'intégration de déshydroalanine. Alors, la fusaristatine C est nécessairement synthétisée via un cluster de gènes différent de celui de la

fusaristatine A dont le premier module de la NRPS est spécifique de l'intégration d'une alanine (Figure 39 C). Cette hypothèse s'applique également dans le cas où la déshydroalanine est issue d'une transformation de l'alanine avant son intégration dans la fusaristatine A (Figure 39 B). D'autre part, il est possible que la transformation de la sérine en déshydroalanine puisse être réalisée après incorporation de la sérine. Alors la NRPS7 possède un domaine responsable de l'incorporation d'une sérine, et la fusaristatine C est également nécessairement synthétisée via une autre NRPS (Figure 39 D).

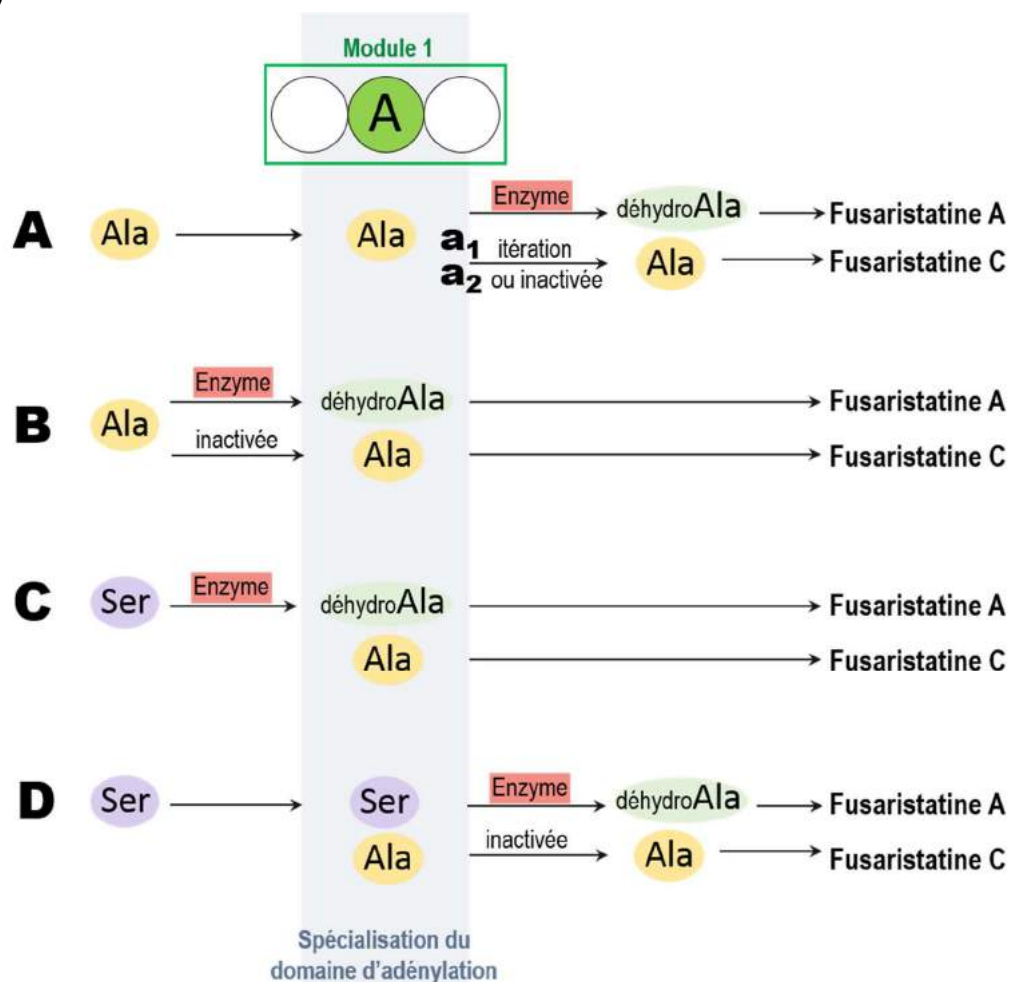


Figure 39. Hypothèses quant à l'origine génétique de la synthèse de fusaristatine C d'après les hypothèses formulées quant à la synthèse de la fusaristatine A par Sorensen et al.<sup>6</sup>.

(3) *Microhétérogénéité permettant au premier domaine d'adénylation d'incorporer une alanine, une sérine ou un déshydroalanine.* Une troisième hypothèse consiste à suggérer que le premier module de la NRPS7 possède un domaine d'adénylation ayant la capacité d'intégrer une alanine, une déshydroalanine ou une sérine. Cette microhétérogénéité a notamment été démontrée dans le cas de la synthèse des fungisporines chez différentes espèces de *Penicillium* et d'*Aspergillus*. Couplée à une inactivation d'enzymes, ou à un phénomène d'itération, cette hypothèse permet d'expliquer les schémas B, C et D (Figure 39) et la synthèse de fusaristatine A et C par le même cluster de gènes. Afin de vérifier la spécificité du premier domaine d'adénylation, une étude de la NRPS7 (GenBank EYB25244.1) a été réalisée avec NRPSpredictor2<sup>8</sup>. Le logiciel a bien repéré la présence de trois domaines d'adénylation, mais les résultats n'ont pas permis de connaître précisément la spécificité du premier domaine d'adénylation.

Des analyses génétiques supplémentaires sont nécessaires afin de rechercher la présence de clusters de gènes similaires à celui responsable de la synthèse de fusaristatine A au sein du génome de *F. graminearum* et qui pourraient potentiellement être impliqués dans la synthèse des nouvelles fusaristatines. De plus, des analyses fonctionnelles sont nécessaires afin de comprendre le rôle de ces métabolites secondaires et d'évaluer leur toxicité. La fusaristatine A a démontré des propriétés inhibitrices de la croissance cellulaire et les deux nouvelles fusaristatines pourraient présenter des activités similaires. En vue de réaliser des tests de toxicité pour comparer les activités de ces métabolites secondaires, une culture de *F. graminearum* à grande échelle et des étapes de purification sont nécessaires. Suite à la purification de la fusaristatine D, des analyses RMN pourraient permettre de caractériser précisément la structure de sa chaîne latérale. La purification de ces molécules est également nécessaire pour réaliser des gammes d'étalonnage et permettre le contrôle de la quantité de ces métabolites dans les denrées alimentaires.

### **2.3. Les métabolomes secondaires des espèces de *Penicillium productrices d'ochratoxine A***

#### **2.3.1. Caractérisation de métabolites secondaires connus**

Lors de l'analyse des métabolomes secondaires de *P. verrucosum* et de *P. nordicum*, certains composés connus qui n'avaient jamais été observés chez ces espèces ont été identifiés. L'aurantiomide C et la LL-P880 $\beta$  ont été détectés chez les deux champignons. Chez *P. verrucosum*, le brévianamide F et l'anacine ont également été mis en évidence pour la première fois et la LL-P880 $\alpha$ , également connue sous le nom de pestalotine, a été observée chez *P. nordicum*. Enfin, huit fungisporines connues ont été caractérisées chez ces champignons.

L'aurantiomide C est un alcaloïde de type quinazoline caractérisé chez *P. aurantiogriseum*. Sa structure est proche de celles des verrucines et de l'anacine qui sont des métabolites connus pour être produits par les espèces de *Penicillium* étudiés dans cette thèse. Tout comme l'anacine, l'aurantiomide C est cytotoxique<sup>9</sup>.

Le brévianamide F, également connu sous le nom *cyclo(L-Trp-L-Pro)*, a démontré une activité inhibitrice contre les bactéries gram positives ainsi que des propriétés antifongiques supérieures à celles de l'amphotéricine B<sup>10</sup>. De plus, des études suggéraient que ce métabolite puisse être utilisé pour le traitement de maladies cardiovasculaires<sup>11</sup> jusqu'à ce que des propriétés hépatotoxiques lui soient conférées<sup>12</sup>. D'autres métabolites secondaires de la classe des brévianamides ont été associés à des propriétés actives telles que des activités nématocides, anti-inflammatoires, ou insecticides<sup>13,14</sup>. Certains brévianamides ayant été détectés chez *P. viridicatum*, il est probable que *P. verrucosum* produise ce type de molécules. L'annotation du brévianamide F dans le cadre de cette thèse n'est donc pas surprenante.

Plusieurs métabolites secondaires de la classe des  $\alpha$ -pyrones, tels que la PC-2 ou la LL-P880 $\gamma$ , sont connus pour être produits par *P. verrucosum* et *P. nordicum*<sup>15</sup>. Bien qu'elles aient déjà été identifiées chez des champignons du genre *Penicillium*<sup>16</sup>, la LL-P880 $\alpha$  et la LL-P880 $\beta$ , aussi connues sous les noms respectifs de pestalotine et hydroxy-pestalotine, n'avaient quant à elles jamais été détectées chez *P. verrucosum* et *P. nordicum*. Ces deux molécules sont faiblement phytotoxiques<sup>17</sup> et ne sont cytotoxiques qu'envers une seule des lignées cellulaires étudiées dans la littérature<sup>18</sup>. Les études fonctionnelles menées sur les métabolites de la famille des pestalotines ont démontré que

selon leur stéréochimie, ces composés pouvaient potentialiser l'effet de la phytohormone gibbérelline A<sub>3</sub><sup>19</sup>.

Au cours de cette thèse, trois métabolites secondaires ont été annotés en tant que tétrapeptides cycliques de la famille des fungisporines (Tableau 8, section 2.3.2 de la discussion). Ces annotations ont été confortées par la détection chez *P. nordicum* d'un gène très similaire au gène *hcpA* responsable de la production de fungisporines chez *P. chrysogenum*. De plus, neuf tétrapeptides linéaires connus ont été identifiés (Tableau 8, section 2.3.2 de la discussion). Bien qu'ils proviennent probablement de l'hydrolyse enzymatique de tétrapeptides cycliques, ces peptides linéaires sont considérés comme des métabolites secondaires à part entière puisqu'ils sont hydrolysés dans le milieu de culture du champignon. Ainsi, la détection de peptides cycliques et linéaires connus a permis de démontrer que *P. verrucosum* est producteur de cinq fungisporines et que *P. nordicum* synthétise huit fungisporines dont quatre en commun avec *P. verrucosum*. Les peptides de séquences IFFV et VFFI n'avaient jamais été observés naturellement mais pourraient provenir des fungisporines connues de structures *cyclo*(FFVI) et *cyclo*(FFIV) respectivement. Aucune étude fonctionnelle n'a été menée sur les propriétés de ces peptides cycliques ou linéaires, mais des études préliminaires ainsi que la similarité de ces métabolites avec d'autres peptides fongiques suggèrent que les fungisporines soient impliquées dans la croissance aérienne du champignon<sup>20</sup>. De plus, de nombreux peptides linéaires fongiques ont démontré des propriétés actives variées (voir introduction, section 5.3), indiquant l'importance de réaliser davantage d'études fonctionnelles quant aux peptides détectés dans cette étude.

La détection de métabolites connus chez *P. verrucosum* et *P. nordicum* permet d'émettre des hypothèses sur les mécanismes de développement de ces champignons et les facteurs impliqués dans leurs toxicités. Ainsi, les fungisporines produites par *P. verrucosum* et *P. nordicum* pourraient être impliquées dans la croissance de ces champignons et donc dans l'infestation des denrées alimentaires. La détection de l'aurantiomide C et du brévianamide F permet quant à elle de suggérer que la cytotoxicité et l'hépatotoxicité des extraits de ces champignons producteurs d'ochratoxine A résultent des effets concomitants de plusieurs métabolites secondaires. Afin de supporter ces hypothèses, les identifications des composés à des niveaux 2 et 3 doivent être vérifiées par purification et analyse RMN. La toxicité des extraits de *P. verrucosum* et *P. nordicum* pourrait également être générée par un ou plusieurs des métabolites secondaires inconnus mis en évidence dans cette thèse.

### 2.3.2. Découverte d'un nouveau peptide cyclique

L'étude des métabolomes secondaires de *P. verrucosum* et *P. nordicum* et la génération de réseaux moléculaires à partir des données obtenues en spectrométrie de masse haute résolution ont permis de mettre en évidence la production de 27 peptides (Tableau 8). Outre les peptides référencés dans la littérature, un nouveau tétrapeptide cyclique produit par *P. verrucosum* et *P. nordicum* a été détecté. L'analyse de son spectre MS/MS a révélé qu'il était composé de deux résidus valine, d'une phénylalanine et suggère que le quatrième résidu corresponde à une tyrosine hydroxylée au niveau de son cycle aromatique. Les mécanismes de fragmentation des peptides cycliques ne permettant pas de déterminer leur séquence à partir de leurs spectres MS/MS, des analyses structurales supplémentaires sont nécessaires à la caractérisation de ce nouveau métabolite secondaire. Sa production en faible quantité par *P. verrucosum* et *P. nordicum* implique la réalisation de cultures à grande échelle et d'étapes de purification afin de pouvoir analyser ce métabolite par MS<sup>n</sup> ou par RMN. Selon l'hypothèse structurale énoncée précédemment, les acides aminés qui composent ce nouveau peptide cyclique

respectent la composition générale des fungisporines connues, c'est-à-dire deux acides aminés aliphatiques consécutifs (valine et/ou isoleucine) et deux acides aminés aromatiques consécutifs (phénylalanine, tyrosine et/ou tryptophane).

Nom du métabolite	Séquence peptidique	Formule Moléculaire	m/z en ESI <sup>+</sup>	m/z en ESI <sup>-</sup>	T <sub>R</sub> (min)	<i>P. verrucosum</i>	<i>P. nordicum</i>
<b>TÉTRAPEPTIDES CYCLIQUES RÉFÉRENCÉS</b>							
Fungisporine C	<i>cyclo</i> (YYVV)	C <sub>28</sub> H <sub>36</sub> N <sub>4</sub> O <sub>6</sub>	525,2708	523,2555	25,8	x	
Fungisporine D	<i>cyclo</i> (FWVV)	C <sub>30</sub> H <sub>37</sub> N <sub>5</sub> O <sub>4</sub>	532,2901	530,2750	36,4		x
	<i>cyclo</i> (YWVV)	C <sub>30</sub> H <sub>37</sub> N <sub>5</sub> O <sub>5</sub>	548,2875	546,2707	29,4		x
<b>TÉTRAPEPTIDES LINÉAIRES RÉFÉRENCÉS</b>							
Fungisporine A hydrolysée	VFFV	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	511,2912	509,2769	18,0	x	x
Potentielle fungisporine B hydrolysée	VYFV	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>6</sub>	527,2864	525,2720	13,0	x	x
Potentielle <i>cyclo</i> (YWVV) hydrolysée	VYVV	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	566,2972	564,2828	12,6	x	x
Potentielle <i>cyclo</i> (YWVV) hydrolysée	WVYV	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>6</sub>	566,2972	564,2824	13,2		x
Potentielle fungisporine D hydrolysée	VFVV	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	550,3015	548,2892	17,4	x	x
Potentielle fungisporine D hydrolysée	WVVF	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	550,3015	548,2892	19,1	x	x
Potentielle <i>cyclo</i> (FFIV) hydrolysée	FIVF	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>5</sub>	525,3046		20,9		x
Potentielle <i>cyclo</i> (YFVI) hydrolysée	IYFV	C <sub>29</sub> H <sub>40</sub> O <sub>6</sub> N <sub>4</sub>	541,2990	539,2860	13,8		x
Potentielle <i>cyclo</i> (YFIV) hydrolysée	VYFI	C <sub>29</sub> H <sub>40</sub> O <sub>6</sub> N <sub>4</sub>	541,2990	539,2860	15,4		x
<b>NOUVEAUX TÉTRAPEPTIDES LINÉAIRES POTENTIELLEMENT ISSUS DE PEPTIDES CYCLIQUES RÉFÉRENCÉS</b>							
Potentielle <i>cyclo</i> (FFVI) hydrolysée	IFFV (ou LVVF)	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>5</sub>	525,3046		19,5		x
Potentielle <i>cyclo</i> (FFIV) hydrolysée	VFFI	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>5</sub>	525,3046		20,2		x
<b>NOUVEAU PEPTIDE CYCLIQUE</b>							
	2 valines 1 phénylalanine 1 tyrosine hydroxylée	C <sub>28</sub> H <sub>36</sub> N <sub>4</sub> O <sub>6</sub>	525,2682	523,2555	25,8	x	x
<b>NOUVEAUX TÉTRAPEPTIDES LINÉAIRES POTENTIELLEMENT ISSUS DE NOUVEAUX PEPTIDES CYCLIQUES</b>							
	VFFV	C <sub>28</sub> H <sub>38</sub> N <sub>4</sub> O <sub>5</sub>	511,2912	509,2769	18,7	x	x
	VFIF (ou VFLF)	C <sub>29</sub> H <sub>40</sub> N <sub>4</sub> O <sub>5</sub>	525,3046		20,9		x
	VIVF (ou VLVF)	C <sub>25</sub> H <sub>40</sub> N <sub>4</sub> O <sub>5</sub>	477,3060		16,9		x
	VR <sub>1</sub> WV (CL <sub>R1</sub> = C <sub>7</sub> H <sub>11</sub> O)	C <sub>30</sub> H <sub>43</sub> N <sub>5</sub> O <sub>6</sub>	592,3102	578,2984	15,5		x
	VR <sub>1</sub> FV (CL <sub>R1</sub> = C <sub>7</sub> H <sub>11</sub> )	C <sub>28</sub> H <sub>42</sub> N <sub>4</sub> O <sub>5</sub>	537,3050	535,2920	23,1		x
	R <sub>1</sub> FFV (CL <sub>R1</sub> = C <sub>4</sub> H <sub>7</sub> O)	C <sub>29</sub> H <sub>38</sub> N <sub>4</sub> O <sub>6</sub>	539,2860	537,2698	26,9		x
	R <sub>1</sub> R <sub>2</sub> VF (CL <sub>R1</sub> +CL <sub>R2</sub> = C <sub>12</sub> H <sub>19</sub> N)	C <sub>30</sub> H <sub>43</sub> N <sub>5</sub> O <sub>5</sub>	576,3165	574,3030	21,0		x
<b>TÉTRAPEPTIDES LINÉAIRES NON CARACTÉRISÉS</b>							
	2 valines 1 phénylalanine 1 tryptophane	C <sub>30</sub> H <sub>39</sub> N <sub>5</sub> O <sub>5</sub>	550,3015	548,2892	18,0	x	x
	?	C <sub>28</sub> H <sub>42</sub> N <sub>4</sub> O <sub>5</sub>	537,3050	535,2920	23,7		x
	?	C <sub>28</sub> H <sub>42</sub> N <sub>4</sub> O <sub>5</sub>	537,3050	535,2920	24,4		x
<b>NOUVEAUX PENTAPEPTIDES LINÉAIRES</b>							
	R <sub>1</sub> GVVF (CL <sub>R1</sub> = C <sub>15</sub> H <sub>11</sub> )	C <sub>38</sub> H <sub>46</sub> N <sub>5</sub> O <sub>6</sub>	668,3415		24,2		x
	R <sub>1</sub> GFFV (CL <sub>R1</sub> = C <sub>11</sub> H <sub>11</sub> )	C <sub>38</sub> H <sub>46</sub> N <sub>5</sub> O <sub>6</sub>	668,3415		23,8		x

Tableau 8. Peptides détectés chez *P. verrucosum* et *P. nordicum*. CL<sub>R1</sub> indique la composition de la chaîne latérale du composé R<sub>1</sub>.

L'analyse *in silico* de la séquence codant pour la protéine HcpA à l'origine de la production des fungisporines chez *P. chrysogenum* a permis de caractériser les quatre domaines d'adénylation qui la composent. Il a ainsi été suggéré que ces domaines aient une sélectivité plutôt large leur permettant la prise en charge de différents acides aminés. Le nouveau composé cyclique pourrait donc être synthétisé via cette même enzyme et l'incorporation de la tyrosine hydroxylée pourrait se faire par l'un des deux domaines d'adénylation impliqués dans la prise en charge des acides aminés aromatiques. Une hypothèse consiste à dire que le peptide cyclique est synthétisé à partir d'acides aminés standards et qu'une enzyme supplémentaire intervient ensuite pour hydroxyliser le peptide sur son résidu tyrosine. Les quelques différences d'acides aminés entre les sites actifs des domaines d'adénylation de la protéine HcpA de *P. chrysogenum* et de la protéine équivalente de *P. nordicum* pourraient expliquer la détection spécifique de ce nouveau peptide chez *P. nordicum* et *P. verrucosum*. Enfin, il se pourrait que le nouveau peptide cyclique soit synthétisé via une enzyme codée par un gène distinct de *hcpA*. Cette enzyme pourrait comporter un domaine d'adénylation particulier permettant

la prise en charge d'une tyrosine hydroxylée ou comporter un domaine responsable de l'hydroxylation d'un résidu tyrosine. Des analyses du génome de *P. verrucosum* et de *P. nordicum* pourraient mettre en évidence la présence d'un deuxième gène similaire à celui responsable de la production des fungisporines.

### 2.3.3. Caractérisation de nouveaux tétrapeptides linéaires

Dix nouveaux tétrapeptides linéaires ont également été détectés au sein des métabolomes secondaires de *P. verrucosum* et *P. nordicum* (Tableau 8). Parmi ces composés, deux ont présenté des séquences suggérant leur production par une enzyme distincte de HcpA. En effet, les peptides VFVF et VFIF (ou VFLF) ne respectent pas la séquence générale des fungisporines, c'est-à-dire deux acides aminés aliphatiques consécutifs puis deux acides aminés aromatiques consécutifs. Ces peptides linéaires pourraient être synthétisés via un gène similaire au gène *hcpA* mais dont la séquence génétique modifierait l'ordre des domaines d'adénylation au sein de l'enzyme correspondante. Ces deux nouveaux tétrapeptides linéaires ayant des séquences similaires, ils sont probablement synthétisés via une même enzyme, dont l'un des domaines d'adénylation présenterait une microhétérogénéité lui permettant d'intégrer à la fois une valine et une isoleucine (ou une leucine).

Un autre nouveau tétrapeptide a démontré une séquence se démarquant de celle des fungisporines. En effet, le composé VIVF (ou VLVF) se compose de trois acides aminés aliphatiques et d'un unique acide aminé aromatique. Il est donc suggéré que celui-ci soit synthétisé via une enzyme distincte des autres. Ce métabolite n'a été détecté que chez *P. nordicum*. La comparaison des génomes de *P. verrucosum* et de *P. nordicum* pourrait alors mettre en évidence un gène similaire au gène *hcpA*, présent uniquement chez *P. nordicum* et qui pourrait coder pour l'enzyme impliquée dans sa synthèse.

Chez *P. nordicum*, deux tétrapeptides linéaires similaires ont été caractérisés en partie et comportent un acide aminé non standard en deuxième position de leurs séquences. Le composé VR<sub>1</sub>WV comporte en deuxième position un résidu avec une chaîne latérale de formule C<sub>7</sub>H<sub>11</sub>O pouvant correspondre à une tyrosine tétrahydrogénée. Le composé VR<sub>1</sub>FV possède quant-à-lui un acide aminé avec une chaîne latérale de formule C<sub>7</sub>H<sub>11</sub> pouvant correspondre à une phénylalanine tétrahydrogénée. Des purifications et des analyses structurales approfondies sont nécessaires à la caractérisation précise de ces deux structures. Néanmoins, les similarités des chaînes latérales du deuxième résidu et des séquences suggèrent que ces deux composés soient issus du même gène. Deux hypothèses sont alors envisageables. Premièrement, ces deux nouveaux tétrapeptides pourraient être issus de l'hydrogénation des deux composés connus VYWV et VFFV par une hydrogénase indépendante. Deuxièmement, ils pourraient être produits par une enzyme codée par une copie du gène *hcpA* associée à un gène codant pour une hydrogénase.

Enfin, deux nouveaux tétrapeptides linéaires ont présenté des séquences avec un ou deux acides aminés non standards et dont les hypothèses structurales étaient plus difficiles à émettre. Le composé R<sub>1</sub>FFV est composé d'un acide aminé N terminal avec une chaîne latérale de formule C<sub>4</sub>H<sub>7</sub>O. Les composés protéinogènes de composition NH<sub>2</sub>+CH+C<sub>4</sub>H<sub>7</sub>O+COOH ont été recherchés dans la littérature. Les structures des composés résultants ont permis de suggérer la présence en position N terminale d'un acide aminé de type tétrahydrofurane ou tétrahydropyrane (ce dernier engageant alors le carbone α dans la formation de son cycle). Les deux acides aminés N terminaux du composé R<sub>1</sub>R<sub>2</sub>VF (la somme des compositions des chaînes latérales de ces deux acides aminés étant C<sub>12</sub>H<sub>19</sub>N)

n'ont pas pu être précisément caractérisés en raison de spectres de masse générés avec un très faible rapport signal/bruit. Des cultures à grande échelle de *P. nordicum* et la purification de ces composés pourraient permettre la génération de spectres de masse de meilleure qualité et la réalisation d'analyses structurales approfondies.

Trois tétrapeptides cycliques n'ont pas pu être caractérisés. Le composé observé à  $m/z$  550.3015 à 18,0 min en ionisation positive chez *P. verrucosum* et *P. nordicum* est un isomère des tétrapeptides connus VFVW et WVVF. L'analyse de son spectre MS/MS a démontré qu'il était composé des mêmes acides aminés, que sa partie N terminale comportait une valine et un tryptophane et que sa partie C terminale comportait une valine et une phénylalanine. Sa séquence précise n'a pas pu être élucidée en raison de la présence de plusieurs ions fragments contradictoires. Puisqu'il se démarque par son temps de rétention, il est probable que ce composé soit différents des deux tétrapeptides linéaires connus cités plus haut. La séquence de l'isomère observé à 18,0 min pourrait donc être WVFW, VVWF ou VVFW. Les deux premières ne respectent pas le type de séquence des fungisporines et suggère la production de ce métabolite par une enzyme distincte. Cette enzyme pourrait alors être la même que celle à l'origine de la production des nouveaux peptides linéaires VFVF et VFIF (ou VFIF). En revanche, si la structure du métabolite observé à 18,0 min est VVFW, alors ce composé pourrait simplement correspondre un peptide linéaire issu de l'hydrolyse enzymatique de la fungisporine D. Enfin, deux métabolites observés à  $m/z$  537,3050 n'ont pas pu être caractérisés. En effet, la présence en quantité bien supérieure du composé VR<sub>1</sub>FV (dont la chaîne latérale de R<sub>1</sub> a pour formule C<sub>7</sub>H<sub>11</sub>) décrit ci-dessus, et à un temps de rétention trop proche de celui de ces deux composés, n'a pas permis d'obtenir des spectres MS/MS interprétables.

Au cours de cette thèse, il a été suggéré que les tétrapeptides linéaires provenaient de l'hydrolyse de tétrapeptides cycliques. En théorie, on rappelle qu'un peptide cyclique de la famille des fungisporine peut être hydrolysé en trois points distincts, engendrant la formation de trois isomères de peptides linéaires. Par conséquent, pour une fungisporine donnée, il serait en théorie possible de détecter d'une part le peptide cyclique, et d'autre part les trois dérivés d'hydrolyse. Or, ni l'une ni l'autre de ces propositions ne sont respectées pour la majorité des nouveaux tétrapeptides linéaires. D'une part, aucun des composés cycliques pouvant en être à l'origine n'a été détecté. D'autre part, à l'exception du peptide observé à  $m/z$  537.3050 présent sous la forme de trois isomères, chaque nouveau peptide linéaire a été détecté sous la forme d'un seul isomère.

Deux hypothèses permettraient d'expliquer ce phénomène. La première hypothèse consiste à dire que ces nouveaux peptides linéaires sont produits par des enzymes codées par des gènes similaires à ceux responsables de la production des fungisporines mais dont le domaine de cyclisation est absent. Dans ce cas, *P. verrucosum* et *P. nordicum* ne produiraient pas les peptides cycliques correspondant à ces nouveaux tétrapeptides linéaires.

La deuxième hypothèse consiste à dire que ces nouveaux tétrapeptides sont issus de l'hydrolyse enzymatique de tétrapeptides cycliques de la même manière que les fungisporines connues mais que les métabolites cycliques d'origine ainsi que les isomères d'hydrolyse ne sont pas détectables par manque de sensibilité de la technique analytique. Cette hypothèse est confortée par l'observation de certaines fungisporines connues dont la présence n'est représentée que par l'intermédiaire de leurs composés linéaires. De plus, certains tétrapeptides connus ont été détectés sous la forme d'un seul isomère d'hydrolyse. C'est le cas notamment des fungisporines A et B qui sont représentées dans cette étude par la présence d'un seul composé linéaire. Cette deuxième hypothèse



suggère donc la découverte de sept nouvelles fungisporines : la *cyclo*(VFVF), la *cyclo*(VFIF) (ou la *cyclo*(VFLF)), la *cyclo*(VIVF) (ou la *cyclo*(VLVF)), la *cyclo*(VR<sub>1</sub>WV) avec la chaîne latérale de R<sub>1</sub> de composition C<sub>7</sub>H<sub>11</sub>O pouvant correspondre à une tyrosine tétrahydrogénée, la *cyclo*(VR<sub>1</sub>FV) avec la chaîne latérale de R<sub>1</sub> de composition C<sub>7</sub>H<sub>11</sub> pouvant correspondre à une phénylalanine tétrahydrogénée, la *cyclo*(R<sub>1</sub>FFV) avec la chaîne latérale de R<sub>1</sub> de composition C<sub>4</sub>H<sub>7</sub>O pouvant correspondre à un tétrahydrofurane ou un tétrahydropyrane, et la *cyclo*(R<sub>1</sub>R<sub>2</sub>VF) avec les deux chaînes latérales de R<sub>1</sub> et R<sub>2</sub> ayant une composition globale de C<sub>12</sub>H<sub>19</sub>N.

#### 2.3.4. Production de nouveaux pentapeptides par *P. nordicum*

L'analyse structurale des deux métabolites secondaires détectés à *m/z* 668,3415 a permis de suggérer la synthèse par *P. nordicum* de deux pentapeptides linéaires. Ainsi, le composé observé à 24,2 min correspondrait à la séquence R<sub>1</sub>GVVF, avec R<sub>1</sub> ayant une chaîne latérale de composition C<sub>15</sub>H<sub>11</sub>. Le composé observé à 23,8 min correspondrait quant-à-lui à la séquence R<sub>1</sub>GFFV, avec R<sub>1</sub> ayant une chaîne latérale de composition C<sub>11</sub>H<sub>11</sub>. En effet, dans les conditions chromatographiques utilisées dans cette étude (colonne en phase inverse, ACN/H<sub>2</sub>O, de 20% d'ACN à 80% d'ACN), le composé R<sub>1</sub>GVVF est observé à un temps de rétention supérieur au composé R<sub>1</sub>GFFV. La composition de la chaîne latérale de R<sub>1</sub>GVVF semble alors compenser et dépasser l'hydrophobicité supérieure d'un résidu phénylalanine par rapport à un résidu valine. Ces deux nouveaux pentapeptides ne sont pas considérés comme étant des potentielles nouvelles fungisporines. Néanmoins, ils pourraient aussi bien être générés par l'hydrolyse de pentapeptides cycliques.

#### 2.3.5. Hypothèses quant aux propriétés des tétrapeptides produits les *Penicillium*

La mutation du gène *hcpA* chez *P. chrysogenum* a démontré que ce gène était impliqué dans le développement du champignon. La mutation induite a été corrélée avec une inhibition de la synthèse des fungisporines et des tétrapeptides linéaires correspondants. Néanmoins, aucune information n'indique précisément quel(s) composé(s) est(sont) à l'origine de la croissance aérienne inhibée lors de la mutation du gène *hcpA*. La totalité, ou certaines de ces fungisporines et/ou de leur dérivés d'hydrolyse pourraient donc être impliqués et des études fonctionnelles sont nécessaires afin de le déterminer. Quoi qu'il en soit, des perspectives sont envisageables pour la limitation du développement des espèces de *Penicillium* pathogènes via l'inhibition du gène impliqué dans la production de ces tétrapeptides. Des études ont démontré que des voies de biosynthèses de métabolites secondaires fongiques pouvaient être inhibées par des composés naturels non toxiques. Par exemple, la synthèse par *A. flavus* de l'aflatoxine B<sub>1</sub>, l'un des composés naturels les plus cancérigènes, peut être inhibée lors d'un contact du champignon avec un extrait de la plante médicinale *Micromeria graeca*<sup>21</sup>. De manière similaire, la découverte d'une molécule non toxique et aux propriétés organoleptiques neutres permettant d'inhiber la synthèse de fungisporines pourrait permettre de freiner la croissance des espèces de *Penicillium* pathogènes. Cela permettrait d'une part la diminution des quantités de denrées alimentaires altérées et donc de limiter les pertes économiques, et d'autre part de limiter indirectement la production de métabolites secondaires toxiques et donc de protéger le consommateur. Le composé inhibant la synthèse des tétrapeptides pourrait par exemple inhiber spécifiquement l'une des enzymes impliquées dans leur synthèse. En conséquence, les substrats de cette enzyme (les intermédiaires de synthèse) s'accumuleront et il est nécessaire que ceux-là ne soient ni toxiques, ni impliqués eux aussi dans la croissance du champignon.

Des études de toxicités sont également nécessaires afin d'évaluer le risque lié à la consommation de fungisporines et autres tétrapeptides. Des études ont démontré que des tétrapeptides cycliques méthylés similaires aux fungisporines n'étaient pas cytotoxiques<sup>22</sup>, mais les nombreux exemples de tétrapeptides fongiques toxiques incitent à vérifier les propriétés des métabolites secondaires identifiés dans le cadre de cette thèse.

Outre l'aspect environnemental, les tétrapeptides détectés chez *P. verrucosum* et *P. nordicum* pourraient présenter un intérêt pharmaceutique. Lors de cette thèse un tétrapeptide de séquence VFVF probablement issu de l'hydrolyse d'un peptide cyclique *cyclo*(VFVF) a été identifié. Ce dernier possède une structure très proche du *cyclo*(N-MePhe-Val-N-MePhe-Val) produit par le champignon *O. sclerotica*<sup>22</sup>. Or ce métabolite secondaire possède des propriétés bloquantes vis-à-vis des canaux calciques cardiaque Cav1.2 (IC<sub>50</sub> 6,2 ± 0,6 μM). Son activité est moins intense que les composés utilisés couramment en thérapeutiques dans la réduction de la tension artérielle (diltiazem IC<sub>50</sub> = 2,2 ± 0,5) mais présente une spécificité vis-à-vis des canaux calciques, avec une inhibition des canaux potassique hERG d'une IC<sub>50</sub> supérieure à 24μM, ce qui présume une absence de cardiotoxicité. Enfin, ce composé méthylé ne présente pas de cytotoxicité ce qui en fait un bon candidat pour une application thérapeutique. Les tétrapeptides découverts dans cette étude pourraient donc présenter des propriétés similaires et être de bons candidats pour le développement de nouveaux médicaments.

### 2.3.6. Perspectives pour l'analyse de ces composés

Dans le cadre de cette thèse, la structure des nouveaux peptides n'a pas pu être décrite précisément en raison de la complexité des extraits fongiques et de la faible quantité de ces métabolites secondaires synthétisée par *P. verrucosum* et *P. nordicum*. Des essais de purification ont été débutés sur trois des fungisporines majoritaires pour leur analyse en RMN. Plusieurs limites ont été rencontrées :

- En vue de la purification de certaines fungisporines, des essais de chromatographie préparative couplée à un détecteur UV et à un système de récolte de différentes fractions ont été développés. La première limite de ces essais de purification s'est trouvée dans le fait que les peptides analysés présentaient une trop faible absorbance quel que soit la longueur d'onde et n'ont pas pu être détectés au sein de l'extrait analysé. Il aurait été possible de situer les peptides étudiés grâce à un encadrement par deux molécules absorbant dans l'UV, mais une deuxième limite a été rencontrée, ne permettant pas de se servir de tels repères.

- En effet, alors que tous les autres composés de l'extrait avaient des profils analytiques stables lors des analyses MS/MS, tous les peptides ont présenté des temps de rétention augmentant progressivement (+ 0,5 min en 10 injections). Aucun article ni expert sollicité n'a permis d'apporter d'explication à ce phénomène. Les spectres MS/MS étant évidemment inchangés et les profils d'élution des différents peptides étant stables l'un par rapport aux autres, il était tout de même possible de suivre spécifiquement chacun des peptides par spectrométrie de masse.

- Afin de poursuivre la purification, un système de collecte a alors été installé sur l'Orbitrap. Lors d'une injection sur colonne analytique, un vingtième du flux sortant a été envoyé vers l'analyseur tandis que le reste de l'extrait a été dirigé vers un système de récolte. Cette méthode a l'avantage de pouvoir suivre les variations d'élution des composés détectables par spectrométrie de masse et

d'adapter en temps réel les temps de récolte de manière très précise (fractions de 0,5 min correspondant précisément à l'élution du composé). Néanmoins, la faible quantité d'extrait étant analysé à chaque injection et l'encrassement des appareils non développés pour des fins préparatives ont été les limites majeures. Malgré cela, les fractions correspondantes, issues des 400 analyses de fractionnement, ont été mélangées et analysées avec un triple quadrupôle afin de s'assurer de leur pureté avant analyse par RMN.

- Après évaporation du solvant et reprise dans les conditions adéquates, une analyse a été réalisée par RMN. Malgré l'apparente pureté des fractions analysées en spectrométrie de masse, les résultats préliminaires ont démontré que chacune des fractions comportait d'autres molécules en quantité bien supérieure à celle du peptide. Une autre des limites rencontrées lors de l'analyse structurale des peptides a donc été la complexité de l'extrait puisque les analyses par RMN ont révélé de nouveaux composés n'ayant pas été ionisés et détectés en spectrométrie de masse.

Face à ces limites, d'autres solutions existent pour l'analyse structurale des nouveaux peptides. Tout d'abord, des cultures à grande échelle de *P. nordicum* sont nécessaires pour obtenir une quantité importante de ces métabolites secondaires. Afin de ne pas encrasser les appareils, une purification préliminaire de l'extrait est conseillée. Enfin, des expériences de fragmentations successives permettraient d'avoir de plus amples informations quant aux groupements non caractérisés des nouveaux peptides. La multiplication des types d'analyses permet d'obtenir des informations complémentaires. Dans le cadre de cette thèse, des fragmentations en mode CID et HCD ont été réalisées. D'autres types d'analyses sont également informatives dans le cadre de séquençage peptidique tels que le mode CID en haute énergie<sup>3</sup>. La dissociation par transfert d'électron (ETD)<sup>23</sup> et la dissociation par capture d'électron (ECD)<sup>24</sup> sont quant-à-elles applicable à des molécules multichargées. Alors que le mode ETD peut être réalisé sur un Orbitrap doté d'une cellule particulière, la CID de haute énergie et l'ECD nécessitent des technologies non disponibles au laboratoire (TOF-TOF et FT-ICR). Ces méthodes d'analyses complémentaires peuvent notamment permettre de différencier les résidus leucines des résidus isoleucine car elles permettent une fragmentation des peptides au niveau des chaînes latérales des acides aminés. En effet, sept des peptides détectés dans cette étude sont composés de l'un de ces deux acides aminés. Bien que la NRPS responsable de la synthèse des fungisporines ne semble incorporer que des isoleucines, rien n'indique que les nouveaux composés ne soient pas composés de leucine. Dans le cadre de cette thèse, l'analyse des composés standards a démontré que les tétrapeptides avec une isoleucine et les même tétrapeptides comportant une leucine pouvaient ne pas être différenciés par leur temps de rétention.

L'élucidation structurale d'un nouveau composé est théoriquement validée par une analyse RMN. Si la purification des composés pour une analyse RMN est trop complexe, des efforts de synthèse de la structure hypothétique et une comparaison de son profil analytique avec la molécule naturelle pourrait néanmoins permettre de valider la structure du nouveau métabolite secondaire. Une autre solution consiste à développer un organisme simplifié capable de produire quasi-exclusivement les métabolites secondaires étudiés. Deux solutions sont alors envisageables. D'une part, le génome de *P. nordicum* pourrait être délété du maximum de clusters non concernés afin que le signal des nouveaux peptides ne soit pas masqué en RMN. D'autre part, une fois identifié, le cluster des nouveaux métabolites secondaires pourrait être transféré chez une espèce au métabolome secondaire restreint. Des technologies telles que CRISPR-Cas 9 permettraient de développer ce type de mutants plus efficacement qu'avec les technologies traditionnelles.

Des outils bioinformatiques ont également été développés pour l'élucidation structurale de peptides. MS-CPA est un programme spécialement conçu pour l'élucidation structurale de peptides cycliques naturels d'origine non ribosomale afin de faire face à la double difficulté de leur analyse<sup>25</sup>. En effet, la séquence des peptides cycliques n'est pas interprétable avec une simple analyse MS/MS et ces molécules naturelles peuvent intégrer des acides aminés non standards. Cet outil a été mis au point pour annoter les spectres de masse en tandem tout en prenant en compte les intensités relatives des ions. Par manque de temps, ce logiciel n'a pas pu être pris en main au cours de cette thèse. Il pourrait être intéressant d'utiliser MS-CPA pour l'élucidation structurale des peptides cycliques détectés dans cette étude. Néanmoins, l'efficacité d'un tel programme pose question. En effet, les outils dédiés à l'élucidation structurale de peptides linéaires ne sont pas parfaitement performants. Il serait donc surprenant qu'un autre logiciel soit plus efficace pour l'élucidation de structures encore plus complexes telles que les peptides cycliques modifiés.

### 3. CLASSIFICATION MÉTABOLIQUE ET NOTION D'ESPÈCE FONGIQUE

---

#### 3.1. Proximité phylogénique et similarité métabolique

*P. verrucosum* et *P. nordicum* sont des champignons très similaires sur le plan morphologique et génétique et dont la séparation phylogénétique semble s'être réalisée tardivement. Dans le cadre de cette thèse, leurs métabolomes secondaires ont été comparés afin d'apprécier la relation entre proximité phylogénique et similarité métabolique. Pour que les métabolomes secondaires soient comparables, les deux champignons ont été étudiés dans les mêmes conditions de culture, traités avec les mêmes protocoles d'extraction et analysés avec les mêmes outils et paramètres analytiques. Les résultats ont montré que ces deux espèces ont des capacités métaboliques similaires dans les conditions d'analyse de cette étude, avec la production de 98 et 92 métabolites secondaires respectivement. Néanmoins, seuls 11 métabolites sont produits en commun par les deux champignons, soit moins de 12% de leurs métabolomes. L'observation d'importantes différences dans les métabolomes secondaires de ces deux espèces voisines suggère que les métabolites secondaires puissent être considérés comme des facteurs discriminants lors de la caractérisation des espèces fongiques tel que le proposent Frisvad *et al.*<sup>26</sup>

#### 3.2. Une souche de *P. verrucosum* chémotype *anacine*

Le faible pourcentage de similarité entre les métabolomes secondaires de *P. verrucosum* et de *P. nordicum* a dans un premier temps souligné la pertinence de la classification des moisissures selon leurs métabolomes secondaires. En introduction nous avons énoncé que trois types de critères de classification complémentaires pouvaient être appliqués pour caractériser un champignon : les critères morphologiques, les critères génétiques et les critères métaboliques. Ce dernier critère a été pris en compte lors de la séparation phylogénique de *P. verrucosum* et de *P. viridicatum*. Pitt *et al.* précisent toutefois que si le métabolome est considéré comme un critère de taxonomie à part entière, il doit se restreindre à la définition de sous-espèces désignées alors comme « chémotypes »<sup>27</sup>. Dans le cadre de cette thèse, la comparaison des métabolomes secondaires de deux espèces fongiques voisines a démontré que de nombreux métabolites peuvent permettre de distinguer deux espèces et suggère que le métabolome secondaire puisse être un critère de classification au-delà des chémotypes.

Néanmoins, l'observation d'un métabolite jugé spécifique de *P. nordicum* dans le métabolome de *P. verrucosum* a remis en question cette idée. En effet, l'anacine est un composé considéré jusqu'à présent comme un métabolite discriminant entre ces deux espèces fongiques. Certains articles de la littérature définissent même *P. verrucosum* comme l'espèce de *Penicillium* productrice d'ochratoxine A et de citrinine mais non productrice d'anacine<sup>28</sup>. Ici, nous avons démontré qu'une espèce de *Penicillium* pouvait être producteur des trois métabolites. Plusieurs questions découlent de cette observation : La production d'anacine par la souche NRRL 5571 remet-elle en question l'identification du champignon en tant que *P. verrucosum* ? Dans ce cas, la production concomitante de citrine rejette également l'identification en tant que *P. nordicum*. De plus, les critères morphologiques et génétiques ont été en accord avec une identification de la souche en tant que *P. verrucosum*. Peut-on alors qualifier une souche de *P. verrucosum* chémotype anacine ? Faut-il sinon considérer une troisième espèce au sein de laquelle les moisissures produiraient à la fois l'anacine et la citrinine ? La réponse à ces questions réside dans le caractère modulable de la définition d'une espèce.

### 3.3. Réflexion autour de la notion de caractérisation métabolique

Ces observations nous ont mené à réfléchir sur le principe d'espèce. Dans cette section est proposée une ébauche de système de classification qui permettrait de considérer le critère métabolique et de dépasser certaines des limites de la classification actuelle. En raison de la forte variation des métabolomes secondaires fongiques selon l'environnement de culture, nous proposons ici que ce critère soit envisagé en second plan après les critères morphologiques et génétiques, pour confirmer une identification ou trancher en cas d'ambiguïté. De plus, nous proposons que le critère métabolique puisse définir une espèce fongique, au-delà des chémotypes, à condition qu'il s'appuie sur un nombre important de métabolites et ne soit pas défini par la négation (tel que « *P. verrucosum* est une espèce **ne** produisant **pas** d'anacine »). Le système de classification proposé ci-dessous repose sur un calcul de score permettant de classer un champignon selon le nombre de métabolites secondaires discriminants partagé avec une espèce définie.

#### 3.3.1. Mise au point du système

Dans un premier temps, le système de classification devrait être mis en place. Pour cela, il est nécessaire de caractériser chimiquement les métabolomes secondaires de toutes les espèces fongiques répertoriées. La première étape consisterait en des études de métabolomique non ciblée réalisées sur différentes souches de chaque espèce. Lors d'une deuxième étape, les données issues des différentes souches d'une même espèce seraient recoupées. Puis, les métabolomes secondaires des différentes espèces seraient comparés entre eux. Ainsi, il serait possible de mettre en évidence lors d'une quatrième étape des métabolites plus ou moins discriminants pour chaque espèce, c'est-à-dire des métabolites secondaires produits par un nombre plus ou moins important d'espèces (Figure 40). Le pouvoir discriminant de chaque composé pourrait être représenté par un indice de discrimination égal à 1 si le métabolite n'a jamais été observé chez une autre espèce, ou inférieur à 1 selon si ce métabolite a été observé plus ou moins fréquemment chez d'autres espèces. Un score pourrait être ensuite accordé à chaque espèce fongique illustrant le nombre de métabolites qu'elle produit et la force de leurs pouvoirs discriminants. Ce score d'une espèce pourrait être calculé en additionnant les indices de discrimination de tous ses métabolites discriminants.

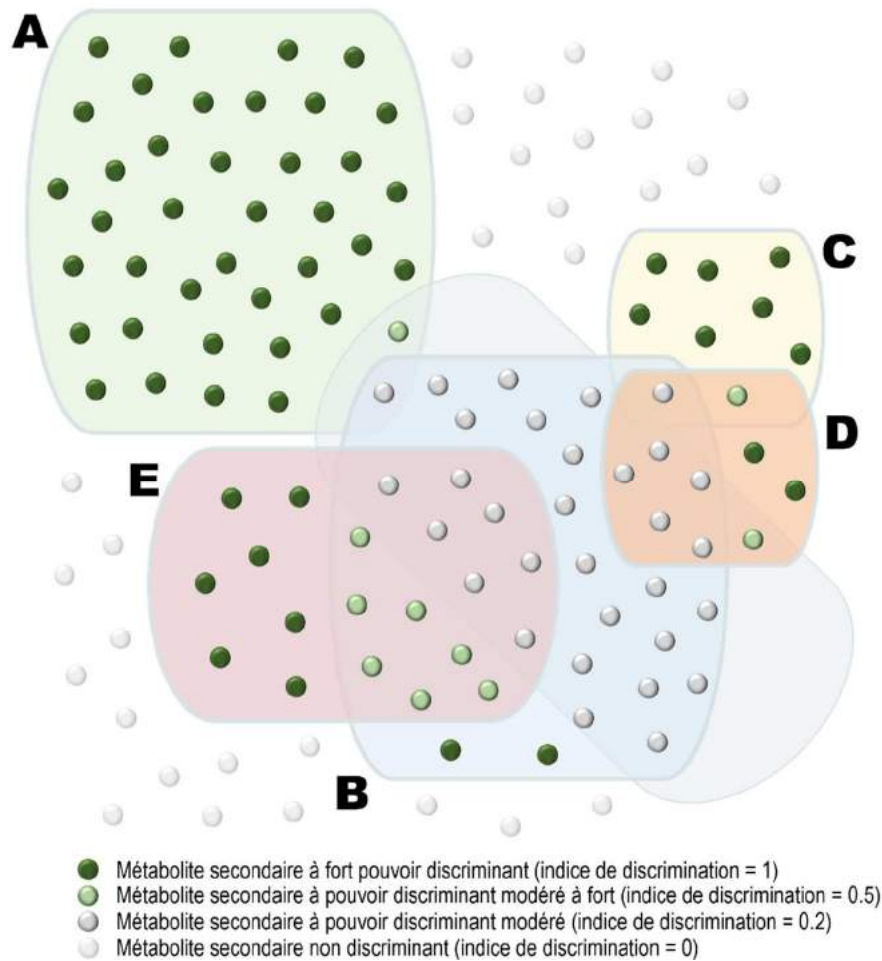


Figure 40. Schématisation d'un système de classification potentiel permettant la classification des espèces fongiques selon leur métabolome secondaire.

La Figure 40 présente une illustration de l'ébauche de système de classification imaginé dans la discussion de cette thèse. Les groupes A à E représentent les groupes de métabolites discriminants de cinq espèces fongiques. Ce schéma prend en considération les différents types d'espèces qui pourraient apparaître : Les espèces avec un grand nombre de métabolites discriminants possédant un fort pouvoir discriminant (Groupe A) ; les espèces ayant un grand nombre de métabolites discriminants mais qui présentent des indices de discrimination plutôt faible (Groupe B), les espèces avec un faible nombre de métabolites discriminants (en raison de leur faible capacité métabolique, ou parce qu'elles produisent en majorité des métabolites non discriminants) mais qui présentent des forts indices de discrimination (Groupe C) ; Les espèces ayant un faible nombre de métabolites discriminants avec des faibles indices de discrimination (Groupe D) ; Et les espèces moyennes (Groupe E). Les scores de chaque espèce pourraient alors calculés comme suit :

$$S_{\text{espèce}} = \sum_{m=1}^n i_m$$

$S_{\text{espèce}}$  étant le score de l'espèce considérée,  $i_m$  étant l'indice de discrimination du métabolite secondaire d'indice  $m$ ,  $m$  allant de 1 à  $n$ , et  $n$  étant le nombre de métabolites discriminants de cette espèce.

Si dans la Figure 40, on considère pour simplifier que les métabolites foncés ont un indice de discrimination égal à 1, que les métabolites verts clairs ont un indice égal à 0,5 et que les métabolites gris ont un indice égal à 0,2 :

- l'espèce A a un score  $S_A = (38 \times 1) + (1 \times 0,5) + (1 \times 0,2) = 38,7$
- l'espèce B a un score  $S_B = (2 \times 1) + (7 \times 0,5) + (31 \times 0,2) = 11,7$
- l'espèce C a un score  $S_C = (7 \times 1) + (1 \times 0,5) + (1 \times 0,2) = 7,7$
- l'espèce D a un score  $S_D = (2 \times 1) + (2 \times 0,5) + (6 \times 0,2) = 4,2$
- l'espèce E a un score  $S_E = (7 \times 1) + (4 \times 0,5) + (7 \times 0,2) = 10,4$

### 3.3.2. Identification d'une souche grâce à ce système

Dans un deuxième temps, une fois le système mis en place, des souches fongiques pourraient être identifiées. Grâce aux critères de classification génétiques et morphologiques, l'identification de ces souches serait tout d'abord orientée. L'analyse du métabolome secondaire de ces souches non caractérisées permettrait ensuite de valider leurs identifications. Par exemple, une souche analysée pourrait être identifiée comme étant de l'espèce A si elle satisfait les critères morphologiques et génétiques de cette espèce, et si elle produit des métabolites secondaires lui permettant d'atteindre un score supérieur ou égal à 80% du score de l'espèce A, c'est à dire supérieur ou égal à 30,96. Cette valeur de score minimum pour être identifié en tant qu'espèce A a été choisie arbitrairement dans cette illustration mais elle devrait être fixée par des experts. Ce seuil pourrait être identique pour toutes les espèces, ou varier selon les espèces.

Le score de la souche analysée est calculé par la somme des indices de discrimination des métabolites secondaires qu'elle produit et qui font partie du groupe de métabolites discriminant de l'espèce A. Il peut donc être constitué de différentes combinaisons de métabolites secondaires, chacun apportant un degré de preuve pour l'identification d'une souche. Ainsi, une souche ayant subi une mutation ponctuelle et ne sécrétant plus l'un des composés discriminant de l'espèce à laquelle elle appartient pourrait tout de même être identifiée comme telle. De plus, une souche ayant acquis la possibilité de produire un métabolite discriminant pour une autre espèce pourra tout de même être caractérisée correctement grâce à la présence d'une majorité de métabolites discriminants référencés chez son espèce. Ce système pourrait donc permettre de prendre en compte la variation du métabolome secondaire des moisissures selon leur environnement et leurs légères modifications acquises au cours d'une évolution continue et non générée par paliers.

### 3.3.3. Analyse critique d'un tel système de classification métabolique

Pour la mise en place d'un tel système, les scores de chaque espèce devraient être calculés après leur culture dans des conditions strictement identiques. Lors de son application, le calcul du score d'une souche à identifier devrait lui aussi être calculé à partir de ces conditions. Il serait donc nécessaire de choisir un milieu de culture et des conditions de cultures standardisées et aisément réalisables (voire commercialisées).

Le développement d'un tel système de classification impliquerait l'analyse non ciblée d'un nombre très important de souches et représenterait un travail considérable. L'utilisation de marquage isotopique est nécessaire pour détecter spécifiquement les métabolites secondaires fongiques et représente une autre limite dans sa création en raison de son coût élevé. Néanmoins, cela contribuerait à l'acquisition de connaissances en terme de métabolomes secondaires fongiques. La caractérisation structurale des métabolites discriminants n'est pas nécessaire, mais chacun d'entre eux doit être clairement identifié par des paramètres analytiques précis (temps de rétention, profils

MS/MS, formule brute). Une autre des limites de ce système est donc la nécessité de générer les données à partir des mêmes protocoles de préparation d'échantillon et des mêmes outils analytiques, à la fois lors de la création du système et lors de l'identification d'une souche.

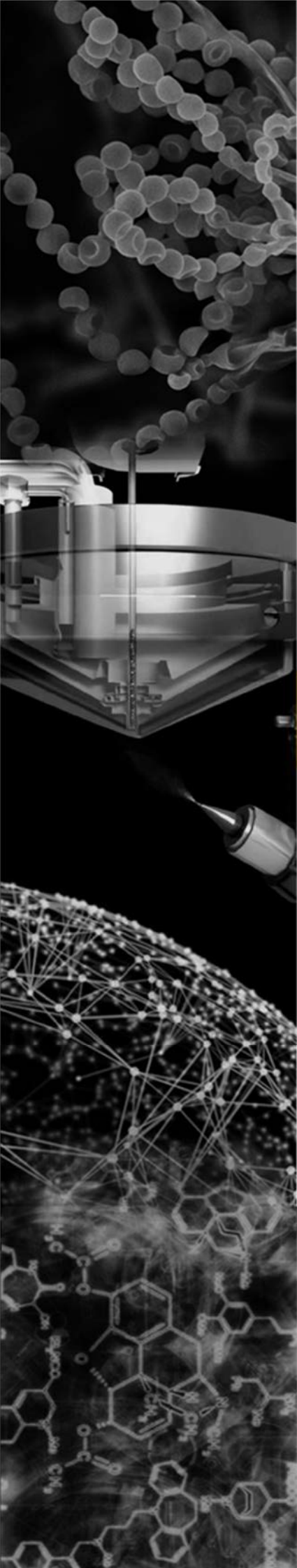
De manière générale, les avancées scientifiques en matière de technologie (capacité de séquençage haut débit, technologies permettant l'analyse globale des métabolomes) font émerger l'idée que le système de classification actuel est loin d'être définitif. La taxonomie a évolué au fil des découvertes scientifiques et des remaniements au sein d'espèces et de genres. De plus en plus de critères sont pris en considération pour la définition d'une espèce et permettent d'éviter les erreurs d'identification. Néanmoins, les travaux menés dans cette thèse suggèrent que la classification des champignons soit un système qui ne devrait pas être fragmenté de manière stricte sur le plan métabolomique. L'espèce caractérisée comme *P. verrucosum* dans cette étude est probablement située à mi-chemin entre le *P. nordicum* classique et le *P. verrucosum* classique en terme de capacité métabolique.



#### 4. RÉFÉRENCES DE LA DISCUSSION

---

- 1 F. Allen, A. Pon, M. Wilson, R. Greiner and D. Wishart, *Nucleic Acids Res.*, 2014, **42**, 1–6.
- 2 I. Blaženović, T. Kind, H. Torbašinović, S. Obrenović, S. S. Mehta, H. Tsugawa, T. Wermuth, N. Schauer, M. Jahn, R. Biedendieck, D. Jahn and O. Fiehn, *J. Cheminform.*, 2017, **9**, 32.
- 3 K. F. Medzihradzky and R. J. Chalkley, *Mass Spectrom. Rev.*, 2015, **34**, 43–63.
- 4 D. Renner and G. Spiteller, *Biomed. Environ. Mass Spectrom.*, 1988, **15**, 75–77.
- 5 T. Lin, A. H. Payne and G. L. Glish, *J. Am. Soc. Mass Spectrom.*, 2001, **12**, 497–504.
- 6 J. L. Sørensen, T. E. Sondergaard, L. Covarelli, P. R. Fuertes, F. T. Hansen, R. J. N. Frandsen, W. Saei, M. B. Lukassen, R. Wimmer, K. F. Nielsen, D. M. Gardiner and H. Giese, *J. Nat. Prod.*, 2014, **77**, 2619–2625.
- 7 D. Tillett, E. Dittmann, M. Erhard, H. Von Döhren, T. Börner and B. A. Neilan, *Chem. Biol.*, 2000, **7**, 753–764.
- 8 M. Röttig, M. H. Medema, K. Blin, T. Weber, C. Rausch and O. Kohlbacher, *Nucleic Acids Res.*, 2011, **39**, 362–367.
- 9 Z. H. Xin, Y. Fang, L. Du, T. Zhu, L. Duan, J. Chen, Q. Q. Gu and W. M. Zhu, *J. Nat. Prod.*, 2007, **70**, 853–855.
- 10 S. Nishanth Kumar, V. S. Nath, R. Pratap Chandran and B. Nambisan, *World J. Microbiol. Biotechnol.*, 2014, **30**, 439–449.
- 11 H. Jamie, G. Kilian, K. Dyason and P. J. Milne, *J. Pharm. Pharmacol.*, 2002, **54**, 1659–65.
- 12 H. Jamie, G. Kilian and P. J. Milne, *Pharmazie*, 2002, **57**, 638–642.
- 13 R. R. M. Paterson, M. J. S. Simmonds, C. Kimmelmeier and A. W. M. Blaney, *Mycol. Res.*, 1990, **94**, 538–542.
- 14 R. B. A. Mehdi, K. A. Shaaban, I. K. Rebai, S. Smaoui, S. Bejar and L. Mellouli, *Nat. Prod. Res.*, 2009, **23**, 1095–1107.
- 15 L. Rahbaek, S. Sperry, J. C. Frisvad and T. O. Larsen, *Biochem. Syst. Ecol.*, 2003, **31**, 313–317.
- 16 M. Geiger, Y. Guitton, M. Vansteelandt, I. Kerzaon, E. Blanchet, T. Robiou du Pont, J. C. Frisvad, P. Hess, Y. F. Pouchus and O. Grovel, *Lett. Appl. Microbiol.*, 2013, **57**, 385–392.
- 17 P. Venkatasubbaiah, C. G. Van Dyke and W. S. Chilton, *Phytochemistry*, 1991, **30**, 1471–1474.
- 18 D. Q. Luo, L. Zhang, B. Z. Shi and X. M. Song, *Molecules*, 2012, **17**, 8554–8560.
- 19 Y. Kimura, A. Suzuki, S. Tamura, K. Mori, M. Oda and M. Matsui, *Plant Cell Physiol.*, 1977, **18**, 1177–1179.
- 20 H. Ali, M. I. Ries, P. P. Lankhorst, R. A. M. Van Der Hoeven, O. L. Schouten, M. Noga, T. Hankemeier, N. N. M. E. Van Peij, R. A. L. Bovenberg, R. J. Vreeken and A. J. M. Driessen, *PLoS One*, 2014, **9**, e98212.
- 21 R. El Khoury, I. Caceres, O. Puel, S. Bailly, A. Atoui, I. P. Oswald, A. El Khoury and J. D. Bailly, *Toxins (Basel)*, 2017, **9**, 87.
- 22 I. Pérez-Victoria, J. Martín, V. González-Menéndez, N. De Pedro, N. El Aouad, F. J. Ortiz-López, J. R. Tormo, G. Platas, F. Vicente, G. F. Bills, O. Genilloud, M. A. Goetz and F. Reyes, *J. Nat. Prod.*, 2012, **75**, 1210–1214.
- 23 J. E. P. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz and D. F. Hunt, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 9528–33.
- 24 R. A. Zubarev, *Mass Spectrom. Rev.*, 2003, **22**, 57–77.
- 25 W.-T. Liu, J. Ng, D. Meluzzi, N. Bandeira, M. Gutierrez, T. L. Simmons, A. W. Schultz, R. G. Lington, B. S. Moore, W. H. Gerwick, P. A. Pevzner and P. C. Dorrestein, *Anal. Chem.*, 2009, **81**, 4200–4209.
- 26 J. C. Frisvad, B. Andersen and U. Thrane, *Mycol. Res.*, 2008, **112**, 231–240.
- 27 J. I. Pitt, *Appl. Environ. Microbiol.*, 1987, **53**, 266–269.
- 28 T. O. Larsen, A. Svendsen and J. Smedsgaard, *Appl. Environ. Microbiol.*, 2001, **67**, 3630–3635.



# Conclusion

*« L'esprit du savant se trouve en quelque sorte  
toujours placé entre deux observations :  
l'une qui sert de point de départ au raisonnement,  
et l'autre qui lui sert de conclusion. »*

*Claude Bernard*



Les travaux de cette thèse ont permis d'investiguer les métabolomes secondaires de trois espèces de moisissures toxigènes. Les résultats ont mis en évidence qu'environ 80% des molécules produites par ces champignons étaient inconnues. Les métabolites secondaires sont des composés non indispensables à la survie d'un organisme, mais synthétisés dans le but de lui conférer un avantage sélectif pour favoriser son développement. Ces produits naturels présentent donc des propriétés actives dont l'Homme peut tirer les bénéfices ou au contraire contre lesquelles il doit se protéger. En accord avec les estimations actuelles, cette thèse a donc mis en évidence que la majeure partie des métabolites secondaires fongiques restait à découvrir. Rapporté au grand nombre d'espèces existantes et aux différents organismes producteurs (champignons, plantes, bactéries...), ces produits naturels représentent donc une niche très importante pour la découverte de molécules bioactives.

La métabolomique est une technologie dont les avancées des dernières décennies en font une stratégie performante pour la recherche de nouveaux produits naturels. Les enjeux de demain ne sont pas tant dans l'augmentation de la sensibilité et de la résolution des analyseurs que dans le développement d'outils bioinformatiques pour le traitement des importants jeux de données qu'ils fournissent. Les logiciels actuellement disponibles assistent la recherche de nouveaux produits naturels et ont l'avantage de raccourcir considérablement le temps d'analyse des données. En revanche, ils présentent des biais qui imposent une vérification systématique des résultats qu'ils fournissent. Par manque d'expertise, par manque de temps ou simplement par excès de confiance en ces outils, les utilisateurs sont parfois amenés à considérer les résultats de ces analyses bioinformatiques comme des vérités établies. Ce manque de recul peut mener à des résultats de recherche erronés et freiner l'avancée des connaissances scientifiques en matière de métabolites secondaires. Une évolution de la performance de ces logiciels d'une part et une prise de recul par leurs utilisateurs d'autre part sont donc nécessaires pour garantir une évolution efficace des découvertes de nouveaux produits naturels.

De nombreux biais observés chez certains logiciels semblent résulter d'un manque d'expertise des développeurs dans l'un des domaines concernés. Il est donc crucial pour l'avancée de la recherche que les expertises soient partagées. La création de pôles interdisciplinaires et collaboratifs à l'échelle d'un laboratoire, nationale ou internationale permettent de potentialiser la recherche, qu'elle soit destinée à la recherche de produits naturels ou d'un point de vue global. Au-delà de la création d'outils, le partage des expertises peut permettre de développer des approches intégrées pour la recherche de nouveaux produits naturels. Notamment, tirer profit des avantages spécifiques aux stratégies génomiques et métabolomiques mais également des méthodes protéomiques et transcriptomiques est un moyen d'investiguer de manière globale les métabolomes secondaires et de mieux en comprendre la composition. De plus, le développement de nouveaux outils tels que CRISPR-Cas9 va certainement révolutionner la recherche dans ce domaine.

Ces dernières décennies, nous assistons à une recrudescence de la recherche concernant les produits naturels. D'une part, la connaissance des mécanismes impliqués dans la toxicité des organismes toxigènes est cruciale pour la maîtrise des dangers qu'ils représentent et pour la protection des populations. D'autre part, cette source importante de molécules actives pourrait permettre de développer de nouveaux médicaments, particulièrement dans le cadre de l'augmentation des résistances aux antibiotiques ou des pathologies émergentes.



## Caractérisation chimique des métabolomes secondaires de *Penicillium* et *Fusarium* par marquage isotopique et spectrométrie de masse haute résolution

**Sous la direction de :** Dr. Isabelle OSWALD et Dr. Emilien JAMIN

---

### Résumé

Les moisissures sont des champignons filamenteux capables de produire de nombreux métabolites secondaires. La caractérisation de ces molécules est devenue un grand défi au cours des dernières décennies en raison de l'émergence de menaces fongiques et de l'intérêt industriel de nombreux produits naturels. Dans ce contexte, une méthode permettant de caractériser l'ensemble du métabolome secondaire de moisissures a été développée et appliquée à la caractérisation des métabolomes secondaires de *Penicillium nordicum*, *Penicillium verrucosum* et *Fusarium graminearum*, trois espèces responsables d'importantes pertes économiques via l'infestation de denrées alimentaires. Le substrat représentant l'unique source de carbone et d'azote des moisissures, chacun des champignons ont été mis en culture sur trois types de grains de blé: (i) grains naturels, (ii) grains marqués à 97% de  $^{13}\text{C}$ , et (iii) grains marqués à 53%  $^{13}\text{C}$  et 97% de  $^{15}\text{N}$ . Les extraits de ces trois cultures ont été analysés par spectrométrie de masse haute résolution. Grâce au marquage isotopique, les métabolites secondaires ont été spécifiquement détectés et leurs formules brutes ont été caractérisées. Une étape de déréplication a ensuite permis l'annotation des molécules connues. La caractérisation de nouveaux métabolites secondaires a ensuite été assistée par la génération de réseaux moléculaires de similarités MS/MS.

L'étude de *P. verrucosum* et *P. nordicum* a permis de détecter 98 et 92 métabolites secondaires respectivement. Parmi eux, 80% étaient inconnus et seulement sept étaient communs aux deux espèces pourtant phylogéniquement très proches. La génération de réseaux moléculaires à partir de ces deux métabolomes secondaires a permis de mettre en évidence un groupe de 25 composés se fragmentant de manière similaire. Seize de ces composés ont été identifiés comme étant des dérivés de fungisporines, des métabolites suspectés d'intervenir dans la croissance aérienne des champignons. Des analyses structurales par MS<sup>n</sup> ont permis d'investiguer la structure des composés inconnus liés aux fungisporines dans le réseau. Ainsi, de nouveaux composés potentiellement impliqués dans l'infestation des denrées alimentaires par *P. verrucosum* et *P. nordicum* ont été caractérisés.

Le marquage du métabolome de *F. graminearum* par des isotopes stables a permis de mettre en évidence la production de 37 métabolites secondaires dont 29 inconnus lorsque le champignon se développe *in vitro*. Les trois quarts de ces composés ont été détectés lors de la croissance de *F. graminearum* dans ses conditions naturelles de croissance, c'est-à-dire lors de son développement sur plante vivante. La réalisation de réseaux moléculaires a permis de détecter deux métabolites secondaires inconnus similaires à la fusaristatine A. Des analyses par MS<sup>n</sup> ont permis d'élucider les structures de ces composés nommés fusaristatine C et D.

### Mots clés

*Penicillium*, *Fusarium*, spectrométrie de masse, marquage isotopique, métabolomique, métabolite secondaire, réseau moléculaire

### Discipline

Pathologie, Toxicologie, Génétique et Nutrition

### Intitulé et adresse du laboratoire

Institut National de la Recherche Agronomique – INRA

Unité Toxalim – UMR 1331

Laboratoire Biosynthèse et Toxicologie des Mycotoxines / Plateforme AXIOM

180, chemin de Tournefeuille – B.P. 93173

31027 Toulouse cedex 9 - France