

## Métabolomique environnementale : forces et faiblesses\*

Jean-Jacques LEGUAY

CEA/Fontenay-Aux-Roses,  
Direction des Sciences du Vivant,  
92 265 Fontenay-Aux-Roses Cedex, France  
<jean-jacques.leguay@cea.fr>

La métabolomique environnementale est l'application de la métabolomique pour caractériser la réponse du métabolome d'un organisme sauvage ou d'élevage à des « stresseurs » naturels (température, lumière, rayonnement ultra-violet...) ou à des « stresseurs » liés à l'activité humaine (toxiques chimiques, produits phytosanitaires, médicaments...). La métabolomique environnementale s'est développée avec un léger retard par rapport aux mêmes approches développées dans le domaine de la recherche académique ou celui de la santé humaine.

Elle permet de caractériser rapidement et simultanément, en utilisant des techniques de *résonance magnétique nucléaire* (RMN) ou des techniques sophistiquées de spectrométrie de masse, des centaines ou des milliers de métabolites. De ce fait, elle diffère substantiellement des méthodes traditionnelles de la biochimie qui ne détecte qu'un ou quelques métabolites. Il en résulte que la métabolomique environnementale est une approche puissante pour découvrir des profils de biomarqueurs correspondant à des réponses biologiques résultant d'exposition à des toxiques environnementaux. Elle permet également de suggérer des pistes pour identifier les voies métaboliques impliquées dans ces réponses et pourrait dans certains cas permettre de déterminer les mécanismes de toxicité de certaines molécules.

La métabolomique permet également de caractériser directement le phénotype d'un individu ou d'une population de n'importe

**Abstract:** *Metabolomic approaches have the potential to contribute to understand how environmental stressors (in broad sens) can affect both human and environmental health. Metabolomic approaches are well documented in human health in general but his powerful in environmental studies are poorly explored. Through four examples, this short review shows real applications of this approach either on animal or plant kingdom. Strengths and weaknesses, concerning particularly the sampling of organisms, the treatment of avalanche of data, the complexity of environmental parameters will be discussed.*

**Key words:** *environmental metabolomics, untargeted metabolomics, contaminants, sampling, phenotypes, genetic*

quel organisme sans avoir besoin de connaître le génotype de cette espèce : ceci est particulièrement précieux, car les espèces sauvages ou d'élevage, sont plutôt des populations que des lignées fixées génétiquement.

À l'aide de quatre exemples tirés de la littérature, des applications de cette approche dans le domaine de pollutions anthropiques terrestres ou aquatiques sur des organismes animaux et végétaux, sont exposées.

À partir de ces exemples, seront mis en évidence les problèmes liés :

- à l'échantillonnage des organismes à étudier dans leur écosystème ;
- à l'analyse de l'avalanche des données issues de la métabolomique ;
- aux traitements statistiques des données ;
- à sa dépendance vis-à-vis des méthodes d'analyse (RMN, spectrométrie de masse).

Enfin les forces et les faiblesses, de cette approche seront analysées.

### L'échantillonnage des organismes à étudier : un point important à considérer

Prendre en compte leur variabilité métabolique est un prérequis non trivial à réaliser. Cette variabilité d'un individu à l'autre dans une population sauvage a plusieurs origines. Elle peut avoir une origine génétique liée aux populations considérées, elle peut dépendre de facteurs locaux tels que température ou accès à la nourriture ou être due à l'âge, au sexe et/ou au stade de développement de l'individu voir à des variables physiologiques comme la maladie. Ces variabilités introduisent un bruit de fond biologique qui peut masquer les différences métaboliques entre individus « stressés » ou en bonne santé. Il est donc nécessaire de se poser la question de la stratégie optimale pour réaliser l'échantillonnage. Ce propos sera illus-

### Abréviations

AD-PMC : Analyse discriminante par régression partielle des moindres carrés (en anglais, PLS-DA : Partial Least Square regression Discriminant Analysis)

PCB : polychlorure de biphényle

CIT : courant ionique total (en anglais, TIC : Total Ion Current)

CG-SM : chromatographie en phase gazeuse couplée à la spectrométrie de masse

CG-SM-TDV : chromatographie gazeuse couplée à un spectromètre de masse à temps de vol (en anglais, CG-TOF-MS, Gaz Chromatography-coupled with a Time Of Flight Mass Spectrometer)

CL-SM-QTDV : chromatographie liquide couplée à un spectromètre de masse quadrupôle à temps de vol (en anglais, LC-QTOF-MS, Liquid Chromatography coupled with a Quadrupole-Time Of Flight Mass Spectrometer)

DL 50 : dose d'un toxique létal pour la moitié de la population traitée

RMN : résonance magnétique nucléaire

\* Conférence issue de la séance du 18 juin 2008 de l'Académie d'Agriculture. Voir la présentation à l'adresse suivante :

[www.academie-agriculture.fr/detail-seance\\_37.html](http://www.academie-agriculture.fr/detail-seance_37.html).

tré par des expériences [1], conduites sur la moule, *Mytilus edulis galloprovincialis*, animal filtreur sédentaire très utilisé pour suivre les atteintes à l'environnement.

Deux questions doivent être posées ; faut-il prélever l'animal du milieu naturel ou faut-il prévoir une phase de stabilisation de l'animal dans un laboratoire dans des conditions contrôlées avant de faire des mesures ? D'un côté, la capture directe mettra en évidence le « vrai » métabolome, mais les variations métaboliques induites peuvent être considérables. Inversement, la stabilisation au laboratoire risque de réduire les variations métaboliques mais aussi d'introduire des variations dues au transport ou aux conditions d'élevage. De plus, les conditions de laboratoire peuvent annuler les effets des « stresseurs anthropogéniques » produits dans l'environnement, ce qui dégraderait l'information recherchée. Pour résoudre ce dilemme, les auteurs [1] ont utilisé la RMN du proton (RMN  $H^1$ ), pour (i) comparer la variation métabolique totale entre des moules prélevées directement dans la mer et celles transportées puis stabilisée au laboratoire et (ii) déterminer la réponse à l'anoxie qui est un stress métabolique naturel. Pour répondre à la question de l'impact sur les mesures de l'âge ou de la taille de la moule de son genre (mâle ou femelle) un certain nombre de paramètres génétiques ou phénotypiques de moules prélevées à Port Quin, UK et à Southampton, UK ont été réalisés. Les moules de même taille sont prélevées à marée basse après être restées 2 heures hors de l'eau tandis que les moules témoins sont maintenues en pleines eaux. Un groupe équivalent est transporté au laboratoire en utilisant un protocole standard [2] puis à leur arrivée elles subissent une séquence de stabilisation avant une expérience d'anoxie. Les mesures de RMN sont effectuées sur des extraits de muscles adducteurs et de manteau. Les spectres RMN révèlent plusieurs classes de métabolites incluant les acides aminés, des osmolytes organiques (ex : glycine, bétaïne...), des molécules phosphorylées et des intermédiaires du cycle de Krebs.

Pour ce qui concerne les analyses réalisées sur les extraits de muscles adducteurs, contrairement à ce qu'attendaient les auteurs, les effets de l'anoxie sont nettement visibles sur les moules récoltées directement en mer alors qu'ils ne sont pas statistiquement différents sur celle acclimatées au laboratoire. Au contraire, les analyses [3] faites sur les extraits de manteau permettent de séparer facilement les réponses métaboliques des moules mâles de celles des moules femelles dans les conditions d'anoxie et de normoxie, aussi bien sur les moules récoltées en mer que sur celles acclimatées au laboratoire (figure 1).

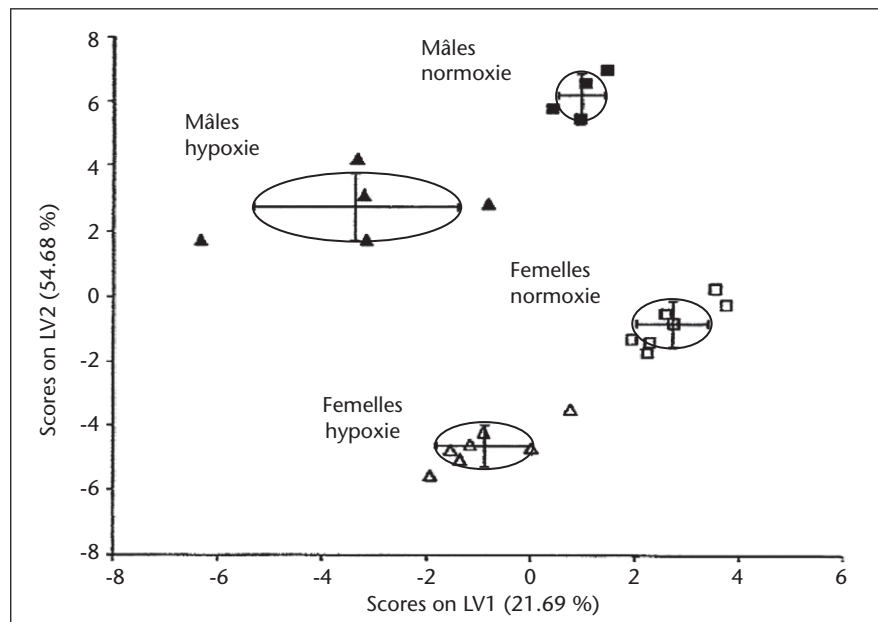


Figure 1. Sont indiqués les scores obtenus selon les deux axes par l'analyse en AD-PMC qui tient compte de la connaissance du sexe des moules. Mâles en hypoxie (triangles pleins), Mâles en normoxie (carrés pleins). Femelles en hypoxie (triangles vides), Femelles en normoxie (carrés vides). L'analyse est réalisée dans un spectromètre à RMN DRX-500 (Brucker Biospin). D'après [1].

L'analyse plus poussée des métabolites (tableau 1) révèle que les différences essentielles sont dues à une concentration plus élevée en glycine, phosphoarginine et glutamate présentes dans les manteaux des moules mâles par rapport à celles du manteau des moules femelles. À partir de ces expériences, il est clair que la stratégie de prélèvement est essentielle pour toute étude environnementale, surtout lorsque l'on mesure le stress métabolique. Deux recommandations sont essentielles. La première est d'effectuer les prélèvements tissulaires et de les fixer immédiatement sur le terrain. La seconde est qu'un certain nombre de traits génétiques (homogénéité de la population) ou phénotypiques (s'assurer du stade de

développement ou du sexage) sont à déterminer avant toute expérimentation.

## Une méthodologie plus performante

Un exemple : la mesure de la toxicité du pyrène sur *Lombricus rubellus*.

Le pyrène est un hydrocarbure polycyclique aromatique formé à partir de combustion incomplète de substances organiques. Certains pyrènes sont cancérigènes [4] d'autres interfèrent avec le système immunitaire et hormonal [5]. Ils persistent dans l'environnement et certains pays se sont dotés d'une législation restrictive pour limiter leur rejet dans l'atmosphère. Les essais

Tableau 1. Importance de la saison de prélèvement : différences dans le métabolisme du manteau de moules mâles et femelles récoltées à Port Quin et à Southampton<sup>a</sup>. D'après [1].

Métabolites	Quin		Southampton	
	Mâle/femelle	P	Mâle/femelle	P
Glycine	4,27	< 0,0001	4,35	< 0,005
Phosphoarginine	3,01	< 0,0001	1,83	< 0,005
Glutamate	1,98	< 0,0001	1,55	< 0,0001
Lysine	0,79	< 0,0500	1,42	< 0,226
Tyrosine	0,65	< 0,0001	1,32	< 0,408
Unidentified <sup>b</sup>	0,65	< 0,0001	0,68	< 0,141
Acetoacetate	0,50	< 0,0001	0,56	< 0,111

<sup>a</sup> Les moules ont été récoltées en un jour en juillet à Port Quin et en 4 fois entre avril et juillet à Southampton. Les rapports des métabolites mâles/femelles sont présentés ensemble avec leur significativité selon le test de Student.

<sup>b</sup> Signal non identifié à 3,68 ppm.

toxicologiques se contentent généralement d'estimer un DL50. L'étude [6] présentée montre que l'analyse des effets sur le métabolome est bien plus performante. Les lombrics sont introduits dans des sols reconstitués contenant des concentrations croissantes de pyrène (0, 10, 40, 160, 640 mg/kg de sol). Après une incubation de 42 jours, les extraits métaboliques des lombrics sont réalisés par un mélange chloroforme-méthanol selon une technique standardisée [7] puis ces extraits sont analysés par RMN et par CG-SM (chromatographie en phase gazeuse couplée à une spectrométrie de masse). La RMN permet de mettre en évidence 32 métabolites différents tandis que la GC-MS, plus puissante en révèle 52. L'observation du comportement des lombrics et leur dénombrement au bout de 42 jours montrent que le test est robuste avec une survie de 100 % pour le témoin, de 90 % pour les 2 plus faibles doses (10 et 40 mg/kg), sans effet sur l'activité de nourrissage ni sur la croissance et une mortalité dose dépendante pour les plus fortes doses (160 et 640 mg/kg) avec une activité moindre de nourrissage et une perte de poids significative. L'analyse métabolomique révèle des désordres métaboliques dès la dose de 40 mg/kg, caractérisés par une forte baisse du lactate et des acides gras insaturés et une forte augmentation de certains acides aminés (alanine, leucine, valine, isoleucine). Ceci suggère fortement que le pyrène entraîne à cette dose une fonte musculaire, les acides aminés étant utilisés pour la néoglucogenèse, ainsi que le lactate.

Ainsi, l'approche métabolomique est non seulement plus sensible pour détecter une perturbation métabolique à une dose 4 fois plus faible que celle du test classique, mais offre de plus des pistes pour expliquer un mécanisme de toxicité du pyrène.

## Comment trouver une « une aiguille dans une botte de foin ? » : La métabolomique permet une recherche non ciblée de contaminants

Même dans le cas de spectres complexes renfermant des centaines de pics une analyse automatisée permet de le faire et permet dans certains cas de repérer, caractériser puis d'identifier une molécule inattendue. Dans cette étude [8], des extraits de huit échantillons de plantes dont deux de plantes témoins sont soumis au test biologique DR CALUX® [9] qui est utilisé pour détecter la dioxine et les molécules de la même famille. Cependant en utilisant une analyse en CG-SM à haute résolution aucun des extraits ne révèle de traces corres-

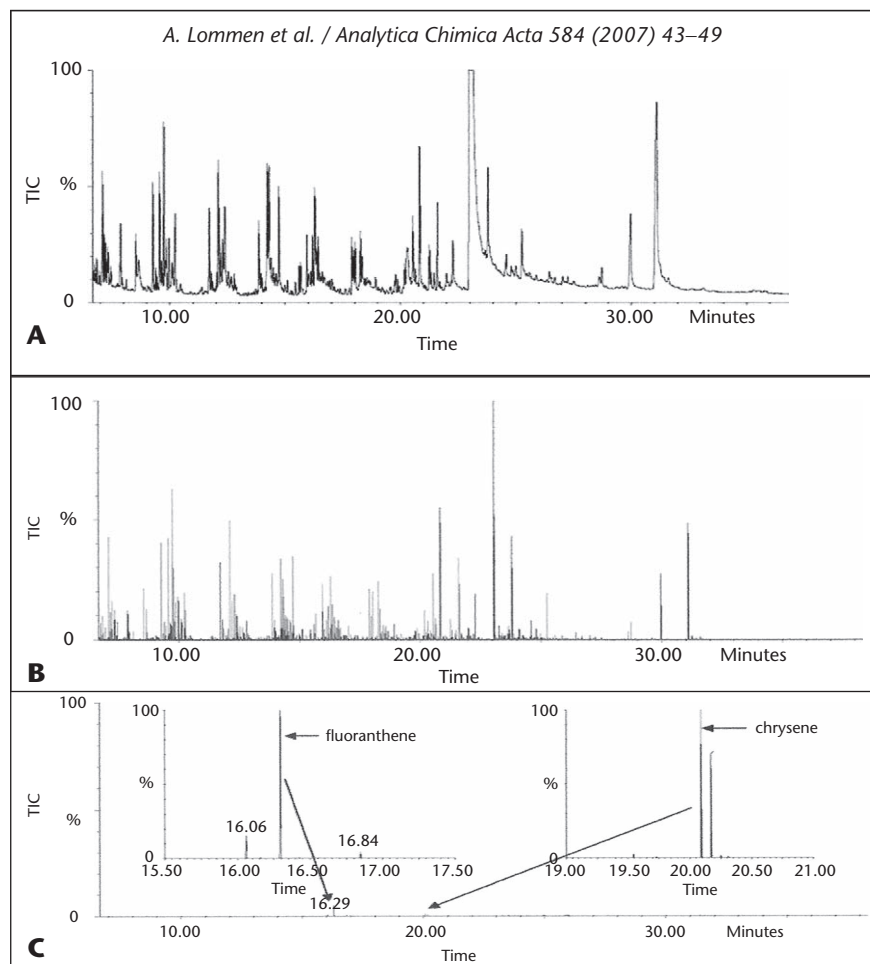


Figure 2. Identification des signaux surnuméraires par CG-SM-TDV présents dans l'extract de plantes répondant positivement au bioessai DR CALUX®. Les échelles verticales sont les mêmes pour les trois profils excepté pour les inserts en (C). L'échelle horizontale est en minutes. A) Courant ionique total (CIT) de la fraction présentant un signal positif au bio essai DR CALUX®. B) Normalisation des signaux (CIT) des résultats de A. C) Signaux sélectionnés (CIT) après soustraction des signaux provenant d'extraits présentant un résultat négatif au bio essai DR CALUX® ; Les inserts agrandis montrent la présence de fluoranthène et de chrysène.

pondant aux 17 congénères de dibenzo-p-dioxine polychlorées et dibenzofurane ni de 12 congénères plans de polychlorure de biphenyle (PCB) ; ce qui jette un doute sur la validité de ce test biologique. L'astuce utilisée par les auteurs [8] consiste simplement à soustraire de la forêt de pics (figure 2) de l'échantillon soupçonné comme « faux positif » les pics des échantillons témoins. Ceci est rendu possible par l'utilisation de metAlign™ [10] qui est un logiciel de traitement automatique de données. Le résultat est spectaculaire puisque seulement 2 pics apparaissent qui sont identifiés respectivement à du fluoranthène et du chrysène qui sont 2 hydrocarbures polycycliques. Le chrysène est effectivement connu [11] pour interférer avec cet essai biologique alors que le fluoranthène n'interfère pas. Ceci montre que le chrysène est probablement responsable du signal positif dans le test DR CALUX® ce qui indique que le protocole standardisé de

traitement des échantillons et qui devrait éliminer le chrysène est insuffisant.

Ce résultat montre bien la puissance de ce type d'analyse et tout laisse supposer que cette approche puisse être étendue à d'autres matrices animales ou végétales.

## Puissance d'analyse et gestion de la complexité

Dans cet exemple [12], il s'agit de montrer l'extraordinaire capacité d'analyse de cette approche pour mettre en évidence, sans *a priori*, des centaines voire des milliers de petites molécules synthétisées par les plantes. La plante sujette à cette étude est *Arabidopsis thaliana*, la plante modèle des laboratoires de génétique et de physiologie végétale encore appelée « arabelle des dames ». Certains pensaient tout connaître de « cette plante de labo-

ratoire » depuis l'obtention du séquençage de son génome puis de l'utilisation de puces à ADN et enfin des études systématiques de l'expression de son protéome. De fait, l'image de son métabolome montre que derrière la simplicité apparente de cette petite plante, se cache une incroyable diversité de molécules. La démonstration est éclatante : les auteurs se sont ingénies à réaliser le profil métabolique de 14 souches sauvages en provenance de différentes régions du globe représentatives de la répartition de l'arabette et de lignées recombinaées descendantes de 2 croisements de lignées pures (*A. Cape verde island* × *A. Landsberg erecta*). Toutes ces lignées ont été autofécondées jusqu'à la 13<sup>e</sup> génération et le taux de région hétérozygote est estimé à 0,71 % à la dixième génération. Toutes ces plantes ont été cultivées dans les mêmes conditions dans des chambres climatiques aux paramètres minutieusement contrôlés. L'analyse métabolomique est réalisée par chromatographie liquide couplé à un spectromètre de masse quadripôle à temps de vol (CL-SM-QTDV).

Les résultats sont les suivants pour les 14 accessions :

– 2 475 métabolites différents identifiés dans les 14 accessions, mais :

- 1 337 métabolites présents pour la plus riche (*A. Cape verde island*)
- 826 métabolites présents pour la plus pauvre (*A. Columbia*)

– 331 métabolites seulement communs aux 14 accessions ;

– 706 métabolites uniques à une accession :

- 235 pour la plus riche (*A. Cape verde island*)
- 14 pour la plus pauvre (*A. Bay-0*)

Pour le croisement (*Cape verde island* × *Landsberg erecta*) l'analyse détecte parmi les 160 lignées, 2 129 métabolites différents dont 853 sont absents des parents.

Ces résultats ne font bien sûr qu'effleurer la diversité métabolique de l'arabette en laissant de côté le reste du monde végétal. Dans cette étude les auteurs ne font le profil que d'une seule population, sous une seule condition de croissance, à un seul stade de développement, avec un seul type d'extraction et avec un seul mode de lecture en spectrométrie de masse. En changeant une seule de ces conditions il est facile d'imaginer que des centaines, voire des milliers de nouveaux pics seraient mis en évidence. En transposant cette étude simplissime aux variations du métabolome sous différentes conditions environnementales dans la nature, on se rend compte que cette approche est encore dans son enfance avant de devenir un outil de diagnostic pertinent.

## Conclusion

Il semble important de lister rapidement les forces et faiblesses de cette approche de métabolomique environnementale en tenant compte des résultats présentés et de ceux qui ne le sont pas.

Du côté de sa force on peut affirmer que :

- les technologies d'analyse sont matures ;
- les profils métaboliques sont intelligibles ;
- les métabolites (petites molécules) sont effectivement la marque ultime des régulations métaboliques ;
- il est possible de définir le normal du pathologique ;
- il est possible de comparer entre les données de terrain et celles du laboratoire

Il est possible de définir des biomarqueurs.

Du côté de ses faiblesses on peut citer :

- un manque de bases de données d'identifications des molécules ;
- un manque de logiciels d'identification automatique et de quantification de données brutes ;
- le fait que des toxiques mineurs en quantité mais important par leurs effets peuvent ne pas être détectés par suite de protocoles d'extraction inadaptés ;
- le fait qu'il est difficile de séparer les causes de la toxicité des conséquences des dommages.

Cette technologie demande encore à être normalisée car il demeure difficile de comparer des jeux de données provenant d'expériences différentes issus de différents laboratoires et obtenus avec des appareils différents. Il serait enfin souhaitable de développer des logiciels intégrés d'analyse métabolomique qui se présenteraient sous la forme de banques de données relationnelles partageables par tous les acteurs de ce domaine. Pour cela, il faudra au préalable trouver un accord au sein de la communauté sur les architectures à créer ainsi que le « vocabulaire et la grammaire » à employer. Cette démarche initiée récemment [13] est indispensable pour imaginer le jour où les agences de réglementation utiliseront ces données en toxicologie humaine, dans un premier temps, puis en toxicologie environnementale ultérieurement.

## RÉFÉRENCES

1. Hines A, Oladiran GS, Bignell JP, Stentford GD, Viant MR. Direct sampling of organisms and knowledge of their phenotype : Key recommendations for environmental metabolomics. *Environ Sci Technol* 2007 ; 41 : 3375-81.
2. Widdows A, Staff FJ. Biological effects of contaminants : Measurement of scope for growth in mussels. *ICES Tech Mar Environ Sci* 2006 ; 40 : 30.
3. Bayne BL. *Marine Mussels : Their Ecology and Physiology*. New York : Cambridge University Press, 1976 ; (506).
4. Baek SD, Field RA, Goldstone ME, Kirk PW, Lester JN, Perr R. A review of atmospheric polycyclic aromatic-hydrocarbons-sources, fate and behavior. *Water Air Soil Pollut* 1991 ; 60 : 287-99.
5. Coles JA, Farley SR, Pipes RK. Effects of fluoranthene on the immunocompetence of the common marine mussel, *Mytilus edulis*. *Aquat Toxicol* 1999 ; 30 : 367-9.
6. Jones OAH, Spurgeon D, Svendsen C, Griffin JL. A metabolomics based approach to assessing the toxicity of the polycyclic aromatic hydrocarbon pyrene to the earthworm *Lumbricus rubellus*. *Chemosphere* 2008 ; 71 : 601-9.
7. Le Belle J, Harris N, Williams S, Bhakoo K. A comparison of cell and tissue extraction techniques using high resolution <sup>1</sup>H NMR spectroscopy. *NMR Biomed* 2002 ; 15 : 37-44.
8. Lommen A, Van Der Weg G, Van Engelen MC, Bor G, Hoogenboom LAP, Nielen MWF. An untargeted metabolomics approach to contaminant analysis : Pinpointing potential unknown compounds. *Analytica Chimica Acta* 2007 ; 584 : 43-9.
9. Gizzi G, Hoogenboom LAP, Von Holst C, Rose M, Anklam E. Determination of dioxins (PCDDs/PCDFs) and PCBs in food and feed using the DR CALUX® bioassay : Results of an international validation study. In : *Food additives and Contaminants : part A ; volume 22*. 2004 : 472-81.
10. The metAlign software can be obtained without costs as download from : <http://www.rikilt.wur.nl/UK/services/MetAlign+download/>.
11. Behnisch PA, Hosoe K, Sakai S. Brominated dioxin-like compounds : in vitro assessment in comparison to classical dioxin-like compounds and other polycyclic aromatic compounds. *Environ Int* 2005 ; 29 : 861-77.
12. Keurentjes JJB, Fu J, Ric De Vos CH, et al. The genetic of plant metabolism. *Nat Genet* 2006 ; 38 : 842-9.
13. Degtyarenko K, De Matos P, Ennis M, et al. ChEBI : a database and ontology for chemical entities of biological interest. *Nucleic Acids Res* 2008 ; 36 : 344-50.