Policy article

Drug Testing and Analysis

Published online in Wiley Online Library

(www.drugtestinganalysis.com) DOI 10.1002/dta.1349

Metabolomics in food analysis: application to the control of forbidden substances

Gaud Dervilly-Pinel,[a] Frédérique Courant,[a] Sylvain Chéreau,[a] Anne-Lise Royer,[a] Fanny Boyard-Kieken,[a,c] Jean-Philippe Antignac,[a,b] Fabrice Monteau[a] and Bruno Le Bizec[a]*

Metabolomics is a science of interest in food analysis to describe and predict properties of food products and processes. It includes the development of analytical methods with the ultimate goal being the identification of so-called 'quality markers', (i.e. sets of metabolites that correlate with, for example, quality, safety, taste, or fragrance of foodstuffs). In turn, these metabolites are influenced by factors as genetic differences of the raw food ingredients (such as animal breed or crop species differences), growth conditions (such as climate, irrigation strategy, or feeding) or production conditions (such as temperature, acidity, or pressure). In cases where the routine-based measurement of a food property faces some limitations such as the lack of knowledge regarding the target compounds to monitor, monitoring based on a limited set of crucial biomarkers is a good alternative, which is of great interest for food safety purposes regarding growth promoting practices. Such an approach may be more efficient than using a classic approach based on a limited set of known metabolites of anabolic compounds. In this context, screening strategies allowing detection of the physiological response resulting from anabolic compound administration are promising approaches to detect their misuse. The global metabolomics workflow implemented for such studies is presented and illustrated through various examples of biological matrices profiling (tissue, blood, urine) and for different classes of anabolic compounds (steroids, β-agonists and somatotropin). © 2012 John Wiley & Sons, Ltd.

Keywords: untargeted profiling; metabolomics; growth promoters; steroids; growth hormone; β-agonists; screening

Introduction

Growth-promoting practices for animal fattening purposes are still encountered all around the world (e.g. use of clenbuterol in pigs, recombinant growth hormone in fish, natural steroids in cocktails, in bovines or hypothetic but realistic anabolic strategies consisting either on upstream disruption of the hypothalamo-pituitary axis (secretagogues, . . .) or even worse on direct genes modification). They are banned in food producing livestock in Europe [1,2] and in other parts of the world. Detection of illegal practices typically relies on residue monitoring in a targeted approach and methods based on gas chromatography (GC) or liquid chromatography (LC) coupled to (tandem) mass spectrometry ((MS/)MS) are today considered as the state-of the-art. [3] These strategies are, however, challenged when facing new xenobiotic growth-promoting agents or new ways of application, such as the administration of low-dose cocktails. To deal with such anabolic practices, alternative means of detecting the abuse are urgently needed. In this context, strategies based on the detection of physiological actions of anabolic practices are promising approaches to screen for their misuse. Profiling biological matrices to reveal biological effects of a drug can either be performed in a targeted focus on a particular class of compounds, such as the steroidome [4–9] or in an untargeted way using global strategies such as transcriptomics, proteomics, or metabolomics. [10,11] These emerging strategies are promising ways to highlight candidate biomarkers to tackle illegal practices. [12] In particular, metabolomics has recently drawn the attention of the scientific community in the detection of anabolic practices and several studies have indeed demonstrated the efficiency of such mass-spectrometric-based fingerprinting to discriminate animals that have been subjected to anabolic treatment from control animals. Several approaches have shown potential to discriminate through analysis of urine samples of animals treated with a range of anabolic formulations: DHEA and pregnenolone in bovines, [13] clenbuterol in calves, [14] recombinant equine growth hormone in horses, [15–17] and nandrolone and estradiol in calves. [18] Metabolomics is an emerging field of 'omics' research that focuses, in an untargeted mode, on large-scale and high-throughput measurement of small molecules (so-called metabolites) in biological matrices. While metabonomics focuses on understanding systemic change through time in complex multicellular systems, [19] metabolomics refers to an analytical description of biological samples, the metabolome consisting in '. . .the complete set of metabolites/low-molecular-weight intermediates, which are context dependent, varying according to the physiology, developmental or pathological state of the cell, tissue,

* *Correspondence to: Prof. Bruno Le Bizec, Oniris, École nationale vétérinaire, agroalimentaire et de l'alimentation Nantes-Atlantique (Oniris), Laboratoire d'Etude des Résidus et Contaminants dans les Aliments (LABERCA), Atlanpole - La Chantrerie, BP 40706, Nantes, F-44307, France. E-mail: laberca@oniris-nantes.fr*

a *LUNAM Université, Oniris, Laboratoire d'Étude des Résidus et Contaminants dans les Aliments (LABERCA), Nantes, France*

b *INRA, Nantes, France*

c *LCH, Verrières le Buisson, France*

organ or organism. . .' [20] The general principle of metabolomics is to characterize biological samples by the production of a chemical profile, i.e. a chemical signature or fingerprint. Until recently, most of the work in metabolomics was focused primarily on clinical or pharmaceutical applications such as drug discovery, drug assessment, clinical toxicology, clinical chemistry, cancer research, or food science and nutrition. [21–24] Within the past few years, metabolomics has also emerged as a field of increasing interest in the context of doping control. If metabolomics only became possible as a result of recent technological breakthroughs in small molecule separation and identification, the complete workflow includes several steps with as many associated skills from the very initial design of the animal experiment to the practical implementation of the tool through sample preparation, sample fingerprinting, data treatment, and analysis and biomarkers structural elucidation. [25]

The present article aims at presenting and illustrating the metabolomics workflow developed and implemented within the French National Reference Laboratory in charge of growth promoters. The relevance and efficiency of this protocol has already been demonstrated through various applications dedicated to the screening of anabolic practices in cattle. [12,14–18] In particular, the specificity related to the different biological matrices that may be used to perform a metabolomics study will be discussed.

Metabolomics workflow to highlight biomarkers of anabolic practices

One of the main limitations in metabolomics analysis compared to other omics approaches is the lack of totally comprehensive approaches, as individuals or laboratories with different skill sets usually develop these. The typical flowchart of a metabolomics study involves several steps (Figure 1), each of them offering a large panel of choices to the scientists, which may have influence on the final result and/or interpretation. As no definitive and unique standard operation procedures currently exist, the comparative generation and analysis of data obtained from identical samples but acquired through different analytical conditions should help the interpretation of the information generated. Such developments relating to analysis and comparisons have recently been conducted in several laboratories, [12–18] leading to the formulation of in-house strategies for each of the steps of interest.

Animal experiment and biological samples of interest

An animal experiment is the compulsory starting point of any metabolomics study aimed at investigating potential biomarkers of anabolic treatment in breeding animals. If the different steps involved in a metabolomics study should all be carefully considered, the design of the experimental protocol is particularly crucial since all subsequent results from the study will rely on the successful achievement of this step. Several parameters have to be determined to ensure collection of valuable samples. It should in particular be pointed out that the conditions such as breed, age, nutrition, and general environment of the animals, may affect the patterns of interest. Therefore, and in order to provide contrasted situations between treated and non-treated animals

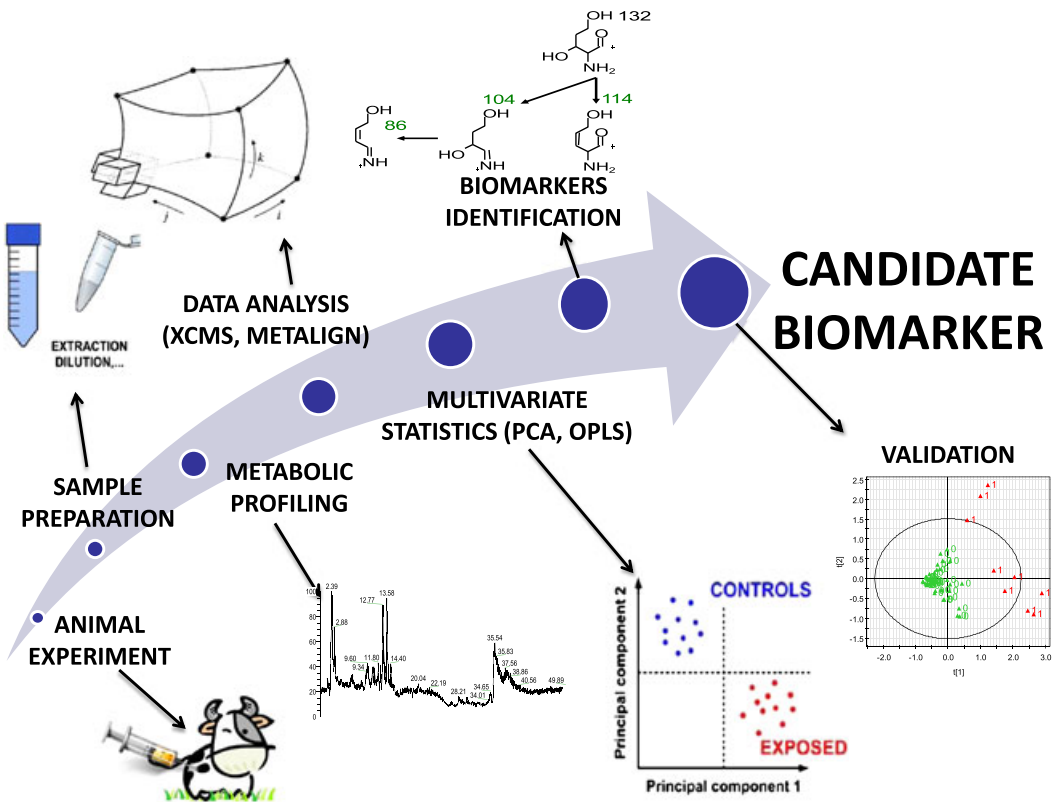


Figure 1. Metabolomics workflow developed and implemented within LABERCA to evidence candidate biomarkers revealing anabolic practices in breeding animals.

which are the more favourable to highlight candidate biomarkers, animal experiments should preferably involve homogeneous animal populations in terms of breed, age, nutrition, and environment.

Because of large inter-individual variations that may occur between individuals involved in the experimental protocol and in order to minimise them, the number of animals as well as the number of resulting samples collected have to be defined to allow sufficient statistical power. Typical reported protocols for the investigation of breeding animals' metabolome to highlight anabolic practices have involved at least ten animals divided into control and treated groups. [14–18]

Regarding biological matrices of interest to be collected during the animal experiment, urine, serum or plasma are easy to collect as often as necessary and can be directly collected at the farm for doping control purposes. In addition, plasma and urine provide very complementary information about the state of an organism. On the one hand, plasma gives an 'instantaneous' readout of the metabolic state at the time of collection, and its composition directly reflects catabolic and anabolic processes occurring in the whole organism. On the other hand, urine provides an 'averaged' pattern of easily excreted polar metabolites discarded from the body as a result of catabolic processes. Such matrices, however, cannot be used to control food of animal origin at the border during importation; therefore, tissues such as muscle or kidney may be more appropriate for such purposes or would fit at the slaughterhouse. Muscle also offers the advantage of being a target tissue for anabolic practices; therefore, the discovery of corresponding biomarkers should be facilitated in this matrix.

Sample preparation

The main drawback of sample preparation in metabolomics is the lack of reference or conventional approaches to be applied. This can be justified by the heterogeneity of applications based on, for example, sample diversity, variability of stimuli or perturbations, and metabolite composition. However, there is a demand for the development of robust strategies for preparation of biological samples to be reproduced in different laboratories. The extremely wide diversity of potential metabolites present in a sample in terms of chemical structures and concentrations makes the goal of measuring them all in metabolomics unrealistic. Nevertheless, the sample preparation procedure has to be as broad as possible in order to avoid losing potentially important information. A compromise is needed between the scope of the analysis *versus* the efficiency and reproducibility of the sample preparation procedure applied. Minimal sample preparation is usually preferred, especially for untargeted applications without any presupposed hypothesis. Conversely, some applications can use relatively selective sample preparation protocols, if some preliminary data or knowledge suggest that the most useful information to look for is present in one or more particular fractions of the sample to be targeted. Typical examples are metabolomics approaches oriented toward very polar and low molecular weight species such as amino acids, [26] or in contrast, towards more lipophilic compounds, [27] that require different analytical procedures.

Liquid matrices (urine, plasma, serum. . .) may be filtered to discard particles and undesirable high molecular weight proteins. A freeze-drying step followed by a reconstitution at a defined dry matter concentration is reported for urine, as a way to avoid the dilution factor issue typically encountered with this matrix. Alternatively, post-acquisition normalization solutions to this problem can be envisaged, for example, normalization of each metabolite signal with the signal intensity of creatine or normalization based on the total signal. [15] The dilution of urine or blood (serum, plasma) samples has also been employed to limit the impact of matrix effects. [28] For solid matrices an extraction step is required for transferring the metabolome compounds into a liquid phase. [29] Samples can be freeze-dried prior to extraction to allow for better homogeneity, repeatability, and extraction capabilities. Different solvent systems with various polarities can be used at this stage (e.g. methanol, acetonitrile, ether, acetone, hexane, cyclohexane) which leads not surprisingly to the production of different and complementary fractions of the studied metabolome. [30] Subsequent purification steps may then be employed (e.g. liquid-liquid partitioning, solid-phase extraction) depending on the degree of selectivity wished to be reached for the particular application. If a gas chromatographic separation is used, the addition of a derivatization reaction step prior to injection may be considered. [31] Metabolic fingerprints acquired after derivatization are usually expected to be more informative since a better chromatographic resolution can be obtained on the less volatile compounds present but at the same time interpretation can be more problematic due to the difficulty to associate a given *m/z* ratio to a derivatized or underivatized compound. In addition, metabolomics requires great care with regard to any potential factor of variability affecting the signals generated, substantially more than for any conventional targeted method and a robust protocol limiting the non-biological variability of the data is highly preferable.

Sample preparation undoubtedly is a crucial stage in metabolomics due to the major issues that have to be kept under control. The sampling conditions appear crucial in terms of representation and homogeneity of the analyzed sample. The sample storage conditions will have a direct influence on the sample stabilisation and integrity. [32] Each step in the sample pretreatment has direct consequences on the metabolic fingerprints generated (content and variability). Overall, the repeatability of this sample preparation process appears to be of prime importance and has to be systematically characterized, considering that each particular choice will directly impact on the obtained metabolic profiles and subsequent interpretation. Most of the applications reported in literature and dealing with anabolic practices detection in breeding animals through metabolomics have been conducted on urine samples and have involved normalization steps through lyophilization. [14,18]

Metabolic profiling

Due to the chemical complexity of the metabolome, it is generally accepted that a single analytical technique will not provide comprehensive visualization of the metabolome, so multiple technologies are generally employed. The selection of the most suitable fingerprinting technology is generally a compromise between speed, chemical selectivity, and instrumental sensitivity. NMR spectroscopy which has traditionally been used in this context presents clear advantages in terms of speed, robustness, and reproducibility. ^{1}H-^{13}C NMR and pyrolysis coupled to metastable atom bombardment (MAB) and time-of-flight (ToF) mass spectrometry had already been used successively to generate metabolic fingerprints from urine samples collected in cattle treated with anabolic steroids, [33,34] but MS-based methods have recently proved to be valuable for such studies, especially thanks to recent technological advances and furthermore present some incomparable advantages over NMR either in terms of sensitivity (Figure 2a). MS measurement following chromatographic

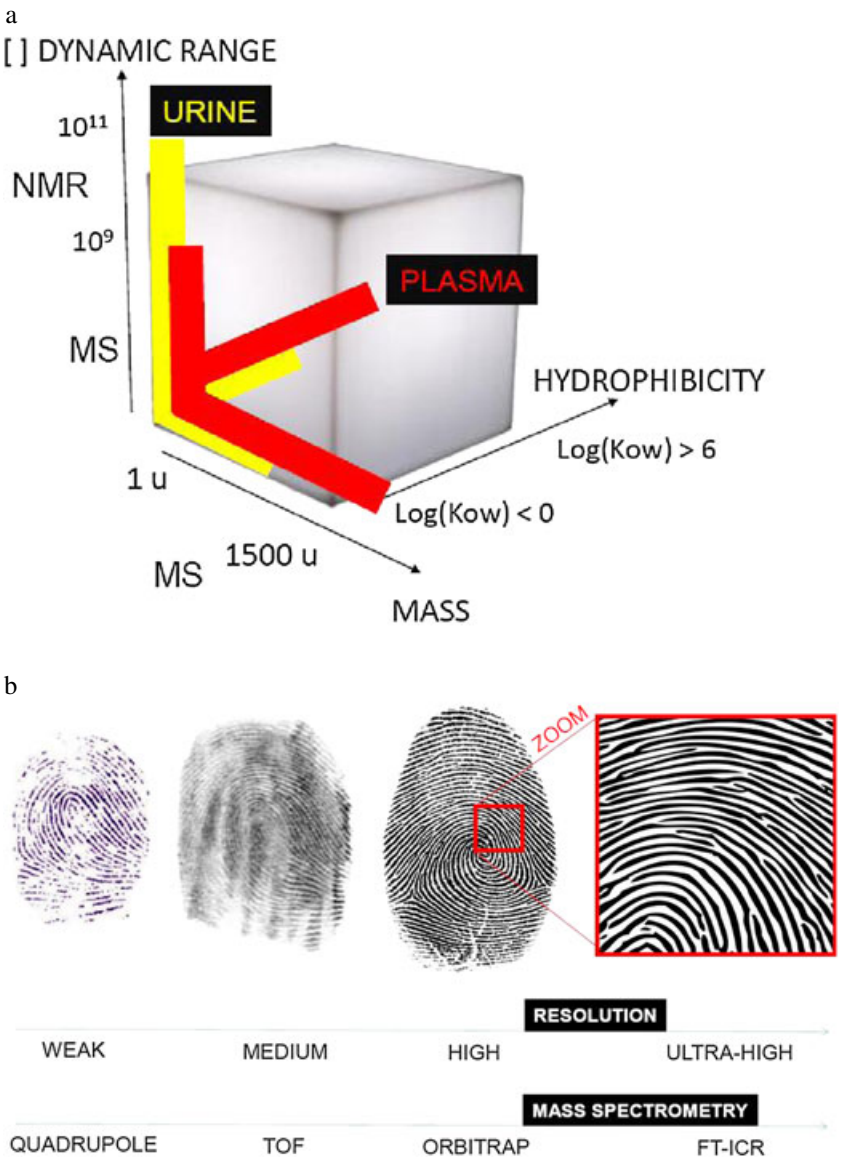


Figure 2. (a) Schematic representation of properties of urinary and blood metabolites. (b) Which MS instrument for which fingerprints precision?

separation offers the best combination of sensitivity and selectivity, so it is central to most metabolomics approaches. Mass-selective detection provides highly specific chemical information including molecular mass and/or characteristic fragment-ion information that is directly related to the chemical structure. This information can be utilized for compound identification through spectral matching with data compiled in libraries for authentic compounds or used for *de novo* structural elucidation. Further, chemically selective MS information can be obtained from extremely small quantities of metabolites with limits of detection in the picomole and femtomole level for many primary and secondary metabolites. Coupling of high resolution separations such as GC, high-performance liquid chromatography (HPLC) or ultra-high-performance liquid chromatography (UHPLC) to MS offers substantial enhancement in metabolome coverage not obtainable through other direct analysis methods using MS or NMR. Based on chemical selectivity, sensitivity, relative cost, and depth-of-coverage, MS has nowadays secured a pinnacle position in metabolomics. Furthermore, MS presents some incomparable advantages over NMR either in terms of capabilities for structural elucidation purposes through its specificity (high resolution and/or MS^n techniques). Indeed, the increasingly high mass-resolving power of modern mass spectrometers such as time-of-flight (ToF), Orbitrap-based technology and Fourier transform platforms allows the simultaneous determination of elemental compositions of thousands of compounds from accurate mass measurements. By combining the advantages of these new analytical tools, high-throughput strategies can be implemented to assess the metabolic profile of biological fluids with high sensitivity and selectivity (Figure 2b). Several applications have been reported in the context of anabolic practices detection in cattle based on ToF MS-finger printings [13,16,18] or Orbitrap. [14,18]

Data processing and analysis

Having acquired data by a measurement method of choice based on MS or NMR, the challenge is to convert these raw, instrumental, data into extracted data (e.g. peak tables) that can easily be processed by statistical tools. The challenges lie in handling issues (e.g. baseline drift, retention-time shifts, noise, and artifacts

generated by the measurement instrument, and the most essential one, detecting the peaks and obtaining a quantitative measurement for each component observed). Detection and quantification of peaks are complicated by the occurrence of overlapping peaks. Preprocessing can be defined as the process of obtaining extracted data (quantified peaks) from raw data, in which these challenges have to be dealt with. Several commercial and open source routines for automatic alignment, smoothing, deconvolution, and extraction of peak have been proposed, such as MS Resolver and ReOrder (Pattern Recognition Systems, Bergen, Norway), [35] MZmine, [36] XCMS, [37] Metalign [38] and MET-IDEA. [39] Many instrument manufacturers have also produced their own software, such as MarkerLynx (Waters), MassHunter (Agilent), MarkerView (AbSCIEX) and SIEVE® (Thermo Fisher Scientific).

Various data-analysis tools are needed to extract the relevant information and several chemometric methods are already extensively applied to find common variation patterns in complex data. Indeed, models based on latent variables such as principal component analysis (PCA) and projection to latent structures (PLS) constitute attractive solutions to provide compact representations of the data and diagnostic tools for the detection of biomarkers. However, as data structures become increasingly complex (multi-way tensors of high order or multi-block data tables), new solutions have been developed to cope with these new types of data. Recently, strong interest has developed in multi-way (PARAFAC, N-PLS) and multi-block methods (CPCA, MB-PLS). A recent review [40] gives an overview of the various strategies that can be applied for metabolomics analysis.

The quality and robustness of the finally obtained models have to be evaluated by different strategies of parameters, such as prediction tests, permutation tests, CV-ANOVA analysis (cross validation-analysis of variance) allowing attributing a degree of significance to the permutation test. In the case of OPLS models, three parameters are of interest for that purpose: $R^2(X)$, corresponding to the proportion of the total variance of the dependant variables that is explained by the model; $R^2(Y)$, defining the proportion of the total variance of the response variable (i.e. the class of the samples) explained by the model; and $Q^2(Y)$, which is similar to $R^2(Y)$ excepted that it is computed by cross-validation. [14–18]

Biomarkers identification

Structural elucidation of the ions of interest is the final and expected step in metabolomics approaches. Indeed, full chemical identification of candidate biomarkers is needed for both extraction of biological sense from the data and, in the objective of screening for growth promoters' administration, the development of efficient protocols based on targeted monitoring of these biomarkers. In particular, it is expected that biological relevance of the identified masses would confirm the suitability of the whole strategy to identify candidate biomarkers. This step, however, remains a challenge. Even if strategies of identification vary amongst the metabolomics community, efforts towards standardization have been realized through the organization of working groups. Dedicated databases have been implemented by various laboratories allowing searching for a compound based on mono-isotopic or molecular mass from accurate mass acquisition, which is unfortunately most of the time insufficient for identification of the compound of interest. Indeed, there are sometimes hundreds of possible structures even for a mass entered with three-decimal-place precision. However, elemental composition may nevertheless be computed from accurate-mass determination and isotopic pattern. Algorithms for filtering molecular formulae have been set up [41] and allow the number of compounds proposed by databases to be reduced. Nevertheless, to use and to benefit from such algorithms, compounds have to be completely resolved from co-eluting or isobaric interferences. Even with the combination of chromatographic separation with MS, this condition is not always fulfilled, so, for structure elucidation of potential biomarkers purposes, the purification and/or concentration (solid-phase extraction, semi-preparative HPLC) of one sample is sometimes needed to isolate the target metabolite and further investigate its identity. The access to multi-dimensional MS is also valuable to help identification with additional structural information. Unequivocal identification requires the combination of complementary techniques. Indeed, as mentioned previously, MS-based techniques can easily provide information. Nevertheless, in the absence of standards to provide the retention time, to add to the MS data, the analyst may still be unable to distinguish between positional isomers. Similarly, LC coupled to 1 H-NMR spectroscopy may provide much of the required structural information but fail in providing a definitive structure in the absence of data derived from MS or IR. This is because there may be present groups (e.g. nitro, nitrile, thiol, phosphate, or sulfate) that do not provide any NMR-detectable signal.

Up till today, only a few biomarkers have been identified and linked to an anabolic treatment in breeding animals. Creatine has been reported in several studies as a compound of interest whose depletion in urine would provide evidence of anabolic treatment with either steroids or β-agonist compounds. [14] One advanced hypothesis is that a less frequently excreted urinary metabolite of this compound would represent a better 'saving' of nitrogen by the organism following this type of anabolic treatment.

Implementation for diagnostic purposes

The main objective of such approaches is both to establish sufficient statistical models to allow prediction of unknown samples and their subsequent classification as compliant or non-compliant (suspect) and also to reveal biomarkers of anabolic practices that could be both used as a basis for screening protocols. In that case, screening methods may be based on targeted and quantitative monitoring of the biomarkers.

Before any implementation, additional experiments involving large numbers of samples from a much broader set of animals are therefore essential to assess the robustness of both models and highlighted candidate biomarkers, irrespective of the surrounding conditions, and to determine their applicability when applied outside the controlled settings of experimental studies. Larger populations of control and treated animals are therefore needed at this stage of the study to come to a robust model suited for routine application.

With regard to the future implementation of these developing and promising strategies in monitoring and surveillance programmes, current EU legislation requires unambiguous evidence of drug administration which can only be achieved through classical targeted MS-based confirmatory approaches. Official implementation of such new screening strategies based on biomarkers monitoring would therefore require modification of the current legislation as recently realized in both human and equine anti-doping codes, where it is, respectively, established that 'Sufficient proof of an anti-doping rule violation is established by either of the following: presence of a prohibited substance or its metabolites or *markers*. . .' [42] and 'The finding of any *scientific*

indicator of administration or other exposure to a prohibited substance is also equivalent to the finding of the substance' [43]. It is clear that the purpose of any profiling strategy should be viewed as an efficient high-throughput screening tool to assist and complement testing programmes utilizing existing confirmatory analytical tools rather than a replacement.

Examples of application

A workflow dedicated to metabolomics studies was implemented some years ago within the French Reference Laboratory for growth promoters. Each step of the process described above has been given particular attention; various options have been tested and optimized to finally implement a more efficient tool (Figure 1) which has already demonstrated its capability in the context of detecting anabolic treatment in cattle. Besides the already published research studies, [14–18] additional examples of application based on the general workflow presented in Figure 1 are described below and in particular, the specificity related to the different biological matrices that may be used to perform a metabolomics study is discussed.

Tissue

While tissue samples have to be collected at the slaughterhouse which may be a drawback compared to biological matrices easy to collect at the farm but also suitable for early screening, they offer the advantage to allow control of food at the border in the case of importation within the EU for instance. As described above, an extraction step is required in the case of solid matrices and the choice of the solvent used is crucial since it determines which part of the metabolome will be extracted and therefore investigated. In order to study metabolic patterns modification in kidney after Revalor S® (140 mg trenbolone acetate, 24 mg 17β-estradiol) treatment, 18 calves (Hereford breed) aged between 6 and 12 months were randomly assigned to the following groups: C (n = 5), T1 (n = 4), T2 (n = 4) and T4 (n = 5) corresponding respectively to control animals, and animals treated with 1, 2, and 4 implants. Kidney samples were collected 90 days after treatment and freeze-dried before analysis. Samples were prepared for metabolomics fingerprinting in order to compare two different extraction solvants – conditions as follows: 100 mg lyophilized tissue was extracted with either methanol or a mixture of methanol/hexane before purification on C18 SPE cartridges eluted with methanol which was further extracted with ether; corresponding dried extract was finally re-suspended in a mixture of acetonitrile/water 20/80 before injection on the LC-ESI(+)-LTQ-Orbitrap™ system. All subsequent steps of the workflow were conducted according to the protocol implemented within LABERCA. Extraction with the mixture composed of methanol and hexane intended to extract at the same time polar and apolar analytes thus giving access to a large metabolome. This extraction condition, however, resulted in non-reproducible fingerprints which could not be further investigated. Data obtained from the analysis of methanolic extracts resulted in fingerprints composed of 2001 ions and allowing the discrimination of the different groups of animals as illustrated on the OPLS displayed in Figure 3, with the characteristics $R^2(X) = 0{,}513$, $R^2(Y) = 0{,}315$ and $Q^2(Y) = 0{,}0982$. A particular focus on ions responsible for the discrimination of the different groups of animals allowed highlighting 2 out of the 2001 constituting the fingerprints as potential candidate biomarkers with significant higher levels in kidney upon Revalor S® treatment (Figure 4a). Further structural elucidation and comparison with standards and in-house database, those ions exhibiting $[M+H]^+$ at $m/z = 450$ and $m/z = 466$ were assigned as chenodeoxycholylglycine and glycocholic acid, respectively, which are both bile products involved in lipids absorption (Figure 4b). It should, however, be noted that based on the OPLS characteristic values and on corresponding permutation test, the robustness of the model could not be fully demonstrated. Mainly, the predictive power of the model linked to $Q^2(Y)$ was not high enough to enable subsequent prediction of

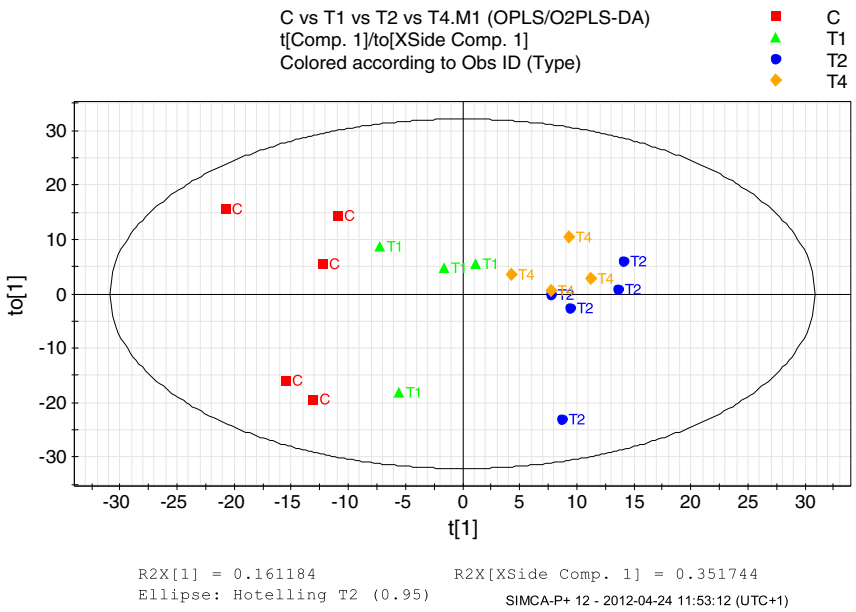


Figure 3. OPLS analysis based on the analysis of kidney after Revalor S® (140 mg trenbolone acetate, 24 mg 17β-estradiol) treatment to 18 calves assigned to the following groups: C, T1, T2 and T4 corresponding respectively to control animals, and animals treated with 1, 2 and 4 implants. Analyses were performed in positive ionization mode with mass range from 50 to 800 *m/z*. The model is based on 2001 ion fragments present with different MS abundances in the methanolic extracts from kidney tissue.

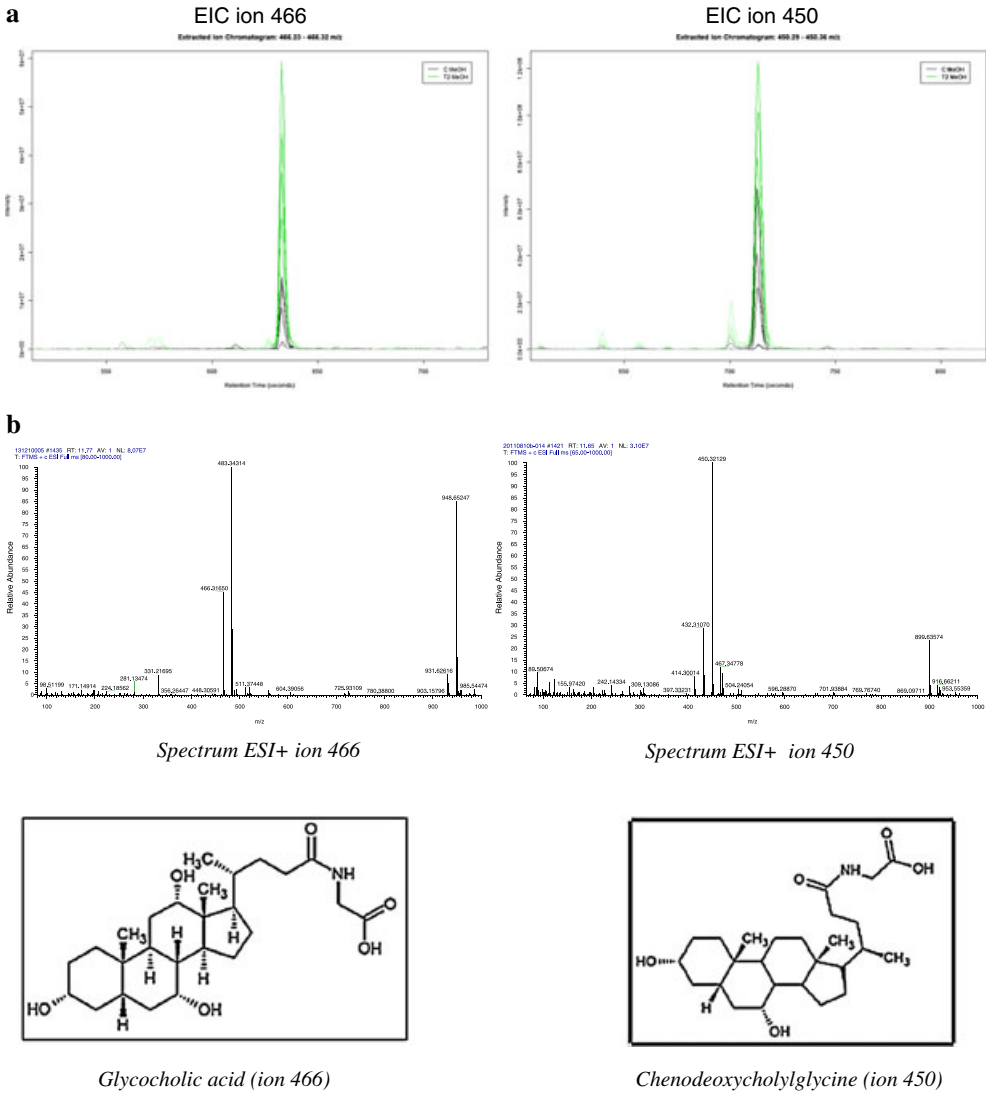


Figure 4. (a) Ion chromatograms of two selected candidate biomarkers showing significant higher abundance in kidney from Revalor S® treated animals than in control animals. (b) Structural elucidation of the two compounds of interest.

unknown samples. Such situation is encountered when the number of variables (n = 2001 (ions) in the present study) is far larger than the number of observations (n = 18 in the present study). These results highlight the importance of the experimental design and the need for sufficient samples for statistical performance requirements. Therefore, and beyond the advantages of tissue as a matrix, the requirement of a large number of samples to establish robust descriptive but mainly predictive models, may constitute one of the main drawbacks of the matrix which is by definition less easy to collect and can only be obtained once from (slaughtered) animals.

Blood

Blood, mainly plasma but also serum, has demonstrated its relevance in the discovery of new biomarkers and has been reported as a matrix of interest in many applications, especially for cancer research purposes and mainly in the frame of proteomic approaches. [44] Although blood is a matrix of interest to tackle anabolic practices through targeted strategies in the field of sport, it is in contrast seldom used to evidence such practices in breeding animals, [45] and even less in the context of untargeted approaches such as metabolomics. [16] Blood is nevertheless an interesting, easy-to-collect biological fluid, which offers the advantage of relative stability with low individual variations. Minimal sample preparation such as filtration of plasma or serum to eliminate high molecular weight protein has been shown enough and efficient to provide informative and reproducible fingerprints.

Metabolomics patterns modifications upon recombinant equine growth hormone (reGH) treatment have been investigated in horse plasma. One stallion and one gelding were daily treated with a pharmaceutical preparation of recombinant equine growth hormone (EquinGen-5®, BresaGen, Australia) over three weeks according to manufacturer's recommendations. Plasma samples (n = 52) were collected before reGH administration (samples C) and during the 30 following days (samples T) and stored at −80 °C until analysis [46]. Plasma samples collected throughout the experiment have been filtered through 10-kDa membrane by centrifugation at 4000 rpm at 6 °C for 90 min and fingerprinted in LC(C_{18})-(ESI+)-HRMS (MicroToF-G-II, Brücker) in the full scan mode (*m/z* in the range 200–1500). All subsequent steps of the workflow were conducted according to the protocol implemented within

LABERCA. Data obtained from the analysis of filtered plasma samples resulted in fingerprints composed of 5086 ions and allowing the discrimination of the different groups of animals as illustrated on the OPLS displayed in Figure 5a, with the characteristics $R^2(Y) = 0.85$ and $Q^2(Y) = 0.50$ evidencing both descriptive and predictive ability of the obtained statistical model. Robustness of the model has been further successfully assessed using permutation test and cross validation test (CV-ANOVA ($p = 0.004$)). Urine samples collected in the frame of the same experiment and previously studied using the same approach also successfully led to the implementation of a robust statistical model allowing differentiation between reGH-treated and non-treated horses. [15] Investigation of potential candidate biomarkers (Figure 5b) in both studies showed, rather unsurprisingly, that none of the ions of interest were common between the two studies. Indeed, both matrices do not give access to similar information in terms of chemical compounds and associated metabolic pathways. Such study highlights the fact that no definitive biological matrix can be defined for a given problem and several of them can be investigated on a complementary basis.

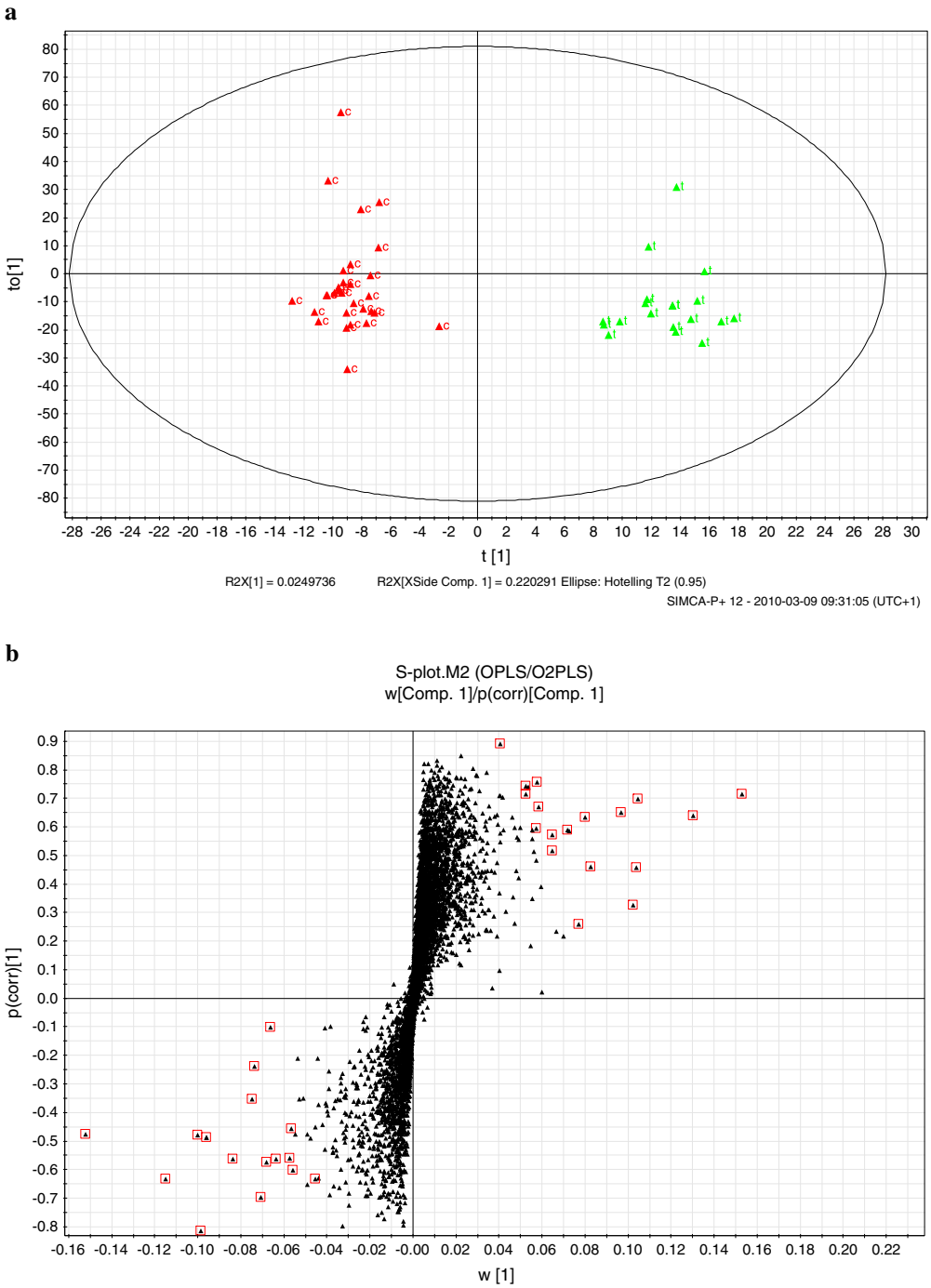


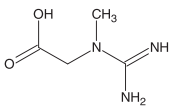
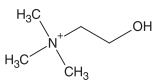
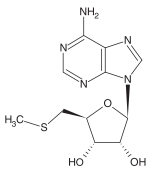
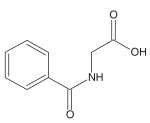
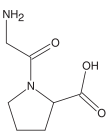
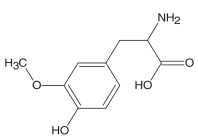
Figure 5. (a) OPLS analyses of plasma samples collected from control and reGH-treated horses. ▲coordinates of samples from control animals; ▲coordinates of samples from GH-treated animals. Analyses were performed in positive ionization mode with mass range from 200 to 1500 *m/z*. The models are based on 5086 ion fragments. (b) S-Plot corresponding to OPLS analysis presented in Figure 5a. Adapted from Kieken *et al.*[16]

Urine

Urine is often reported as a matrix of interest since it is a fluid of excretion containing a large range of potential interesting metabolites which can easily be collected. A number of metabolomics studies proving the relevance of the approach to discriminate situations involving anabolic-treated and non-treated animals have already been reported using urine as the matrix of choice. [13–15] Sample preparation generally involves a filtration followed by a normalization step to correct inter- and intra-individual variations of urine's concentration and allow generation of comparable fingerprints. Significant advances have been obtained in that context and the proof of concept is now widely accepted by the scientific community. The next important step for research laboratories consists of the implementation of the corresponding developed screening tool. This involves elucidation and identification of highlighted candidate biomarkers in a first step followed by their targeted monitoring in a second step.

Urine samples collected in the frame of a previous large-scale animal experiment (n = 24 calves, both male and female) involving treatment with a mixture of 17β-oestradiol (17β-E2) benzoate and 17β-nortestosterone (17β-NT) laureate have been investigated through LC-HRMS fingerprintings which successfully enabled differentiating treated and non-treated animals on the basis of their metabolomics profiles. Details of the animal experiment and fingerprinting research work have been published previously. [4,18] This work was recently continued and a study focusing on the main ions responsible for the discrimination of the animals led to the identification of six of them, namely creatine, choline, 5-methyl-thio adenosine, hippuric acid, methoxy-tyrosine and the dipeptide Gly-Pro (Table 1), which are all small metabolites (Mw < 300 g/mol) widespread in the general metabolism pathways. Focusing on these six candidate biomarkers, a statistical model was built with efficient associated descriptive performances for the discrimination of 17β-E2/17β-NT-treated and control calves (Figure 6a). As a first assessment of the model's robustness, good prediction of additional urine samples collected during the same animal experiment was obtained (Figure 6b). An important step forward was the efficient prediction of urine samples collected in the frame of a different

Table 1. Chemical compounds highlighted in urine as candidate biomarkers of 17β-oestradiol (17b-E2) benzoate /17β-nortestosterone (17β-NT) laureate administration to calves.

CAS	Name	Structure	Formula	M_{mi} (DA)
57-00-1	Creatine		C4H9N3O2	131.06948
67-48-1	Choline		C5H14NO	104.10699
2457-80-9	5-methyl-thio-adenosine(MTA)		C5H15N5O3S	297.08956
495-69-2	Hippuric Acid		C9H9NO3	179.05824
704-15-4	Glycylproline (Gly-Pro)		C7H12N2O3	172.08479
300-48-1	Methoxy-l-tyrosine		C10H13NO4	211.08446

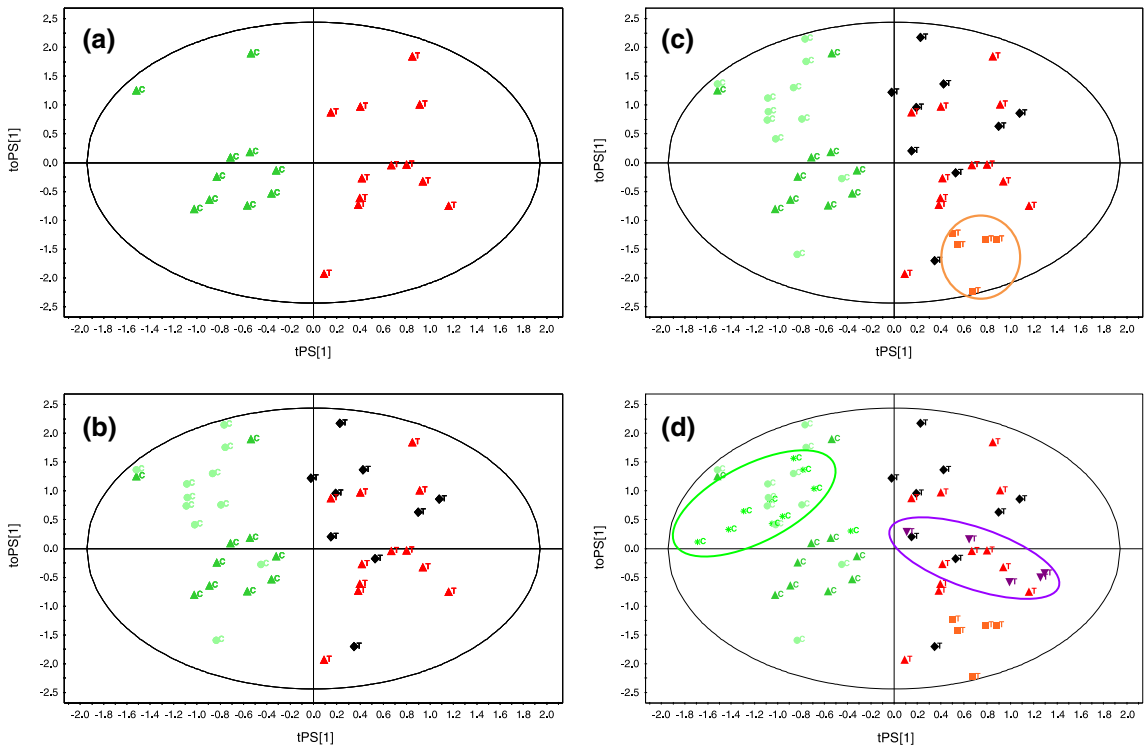


Figure 6. (a) OPLS analyses of urine samples from control ▲ and (17β-E2/17β-NT) treated ▲ calves. (b) Prediction of different urine samples from control ● and (17β-E2/17β-NT) treated ▲ calves. (c) Prediction of urine samples from 17β-E2-treated steers ■. (d) Prediction of urine samples from control ★ and clenbuterol treated ▼ calves.

animal experiment (adult castrated males (steers), 17β-oestradiol benzoate treatment only) (Figure 6c). Such performance demonstrated the ability of the model to take into account the animals' variability (sex, age, environment, and feed) to release information only related to steroid administration, which was the expected main initial objective of the approach, and was never demonstrated before. Finally, the prediction of urine samples collected in the frame of a totally different doping scheme, i.e. clenbuterol administration to calves, resulted in excellent classification of both treated and control samples (Figure 6d). Such very promising results evidence the robustness of the developed model to deal with bovines' variability (genetic, breeding conditions. . .) and also to encompass different treatments such as steroids and β-agonists. It furthermore tends to conclude on the ability of the model to highlight global anabolic effects, irrespective from the type of treatment (steroids, β-agonists. . .) compound. Therefore the urinary levels of the six highlighted candidate biomarkers, when used in balanced combination, can be considered as a promising screening tool in detecting anabolic practices in bovines. Further work is currently ongoing, on the one hand to increase the model's robustness with the inclusion of a large number of control samples, on the other hand to assess specificity of the model towards other classes of anabolic compounds and other animal species.

Conclusion

The need for new and alternative analytical strategies to highlight illegal anabolic practices has caught the attention of scientists over the past decade. Metabolomics in particular was investigated in the hope of identifying biomarkers to provide evidence for such animal breeding practices. Since then, several studies have demonstrated the effectiveness of using MS-based fingerprinting techniques to discriminate control animals from animals treated with a range of anabolic formulations. These types of research, conducted under well-controlled experimental conditions, were necessary steps in the initial assessment of the strategy. While steps towards implementation of corresponding screening tools are still numerous, further consideration will have to be given to the validation of these protocols in order to permit the inclusion of such strategies in the regulation.

Conflicts of interest

The authors have no conflicts of interest to declare.

References

[1] European Commission. Council Directive 88/146/EC of 7 March 1988 prohibiting the use in livestock farming of certain substances having a hormonal action. *Off. J. Eur. Union* **1988**, *L070*, 16.
[2] European Commission. Council Directive 96/23/EC of 29 April 1996 on measures to monitor certain substances and residues thereof in live animals and animal products and repealing Directives 85/358/EEC and 86/469/EEC and Decision 89/187/EEC and 91/664/EEC. *Off. J. Eur. Union* **1996**, *L125*, 10.
[3] B. Le Bizec, G. Pinel, J.-P. Antignac. Review: Options for veterinary drug analysis using mass spectrometry. *J. Chromatogr. A* **2009**, *1216*, 8016.
[4] G. Pinel, L. Rambaud, G. Cacciatore, A. Bergwerff, C.T. Elliott, M.F.W. Nielen, *et al.* Elimination kinetic of 17ß-estradiol 3-benzoate and 17ß-nandrolone laureate esters metabolites in calves urine. *J. Steroid Biochem.* **2008**, *110*, 30.
[5] G. Pinel, L. Rambaud, F. Monteau, C.T. Elliott, B. Le Bizec. Estranediols profiling in calves' urine after 17β-nandrolone laureate ester administration. *J. Steroid Biochem.* **2010**, *121*, 626.
[6] S. Anizan, E. Bichon, F. Monteau, N. Cesbron, J.P. Antignac, B. Le Bizec. A new reliable sample preparation for high throughput

focused steroid profiling by gas chromatography–mass spectrometry. *J. Chromatogr. A* **2010**, *1217*, 6652.
[7] G. Dervilly-Pinel, L. Rambaud, P. Sittishack, F. Monteau, A. Hewitt, G. Kennedy, *et al*. 5α-estrane-3β,17β-diol and 5β-estrane-3α,17β-diol: definitive screening biomarkers to sign nandrolone abuse in cattle? *J. Steroid Biochem.* **2011**, *126*, 65.
[8] S. Anizan, E. Bichon, D. Di Nardo, F. Monteau, N. Cesbron, J.P. Antignac, *et al*. Screening of androstenedione misuse in cattle by LC-MS/MS profiling of glucuronide and sulfate steroids in urine. *Talanta* **2011**, *86*, 186.
[9] J. Boccard, F. Badoud, E. Grata, S. Ouertani, M. Hanafi, G. Mazerolles, *et al*. A steroidomic approach for biomarkers discovery in doping control. *Forensic Sci. Int.* **2011**, DOI: 10.1016/j.forsciint.2011.07.023
[10] I. Riedmaier, C. Becker, M.W. Pfaffl, H.H.D. Meyer. The use of omic technologies for biomarker development to trace functions of anabolic agents. *J. Chromatogr. A* **2009**, *1216*(46), 8192.
[11] C. Reichel. OMICS-strategies and methods in the fight against doping. *Forensic Sci. Int.* **2011**, DOI: 10.1016/j.forsciint.2011.07.031
[12] G. Pinel, M. Mooney, S. Weigel, J.-P. Antignac, M.W.F. Nielen, C. Elliott *et al*. Targeted and untargeted profiling of biological fluids to screen for anabolic practices in cattle. *Trends Anal. Chem.* **2010**, *29*, 1269.
[13] J.C.W. Rijk, A. Lommen, M.L. Essers, M.J. Groot, J.W. Van Hende, T.G. Doeswijk, *et al*. Metabolomics approach to anabolic steroid urine profiling of bovines treated with prohormones. *Anal. Chem.* **2009**, *81*, 6879.
[14] F. Courant, G. Pinel, E. Bichon, F. Monteau, J.-P. Antignac, B. Le Bizec. Development of a metabolomic approach based on liquid chromatography-high resolution mass spectrometry to screen for a clenbuterol abuse in calf. *Analyst* **2009**, *134*, 1637.
[15] F. Kieken, G. Pinel, J.P. Antignac, F. Courant, F. Monteau, M.A. Popot, *et al*. Development of a metabolomic approach based on LC-ESI-HRMS measurement for profiling of metabolic changes induced by recombinant equine growth hormone in horse urine. *Anal. Bioanal. Chem.* **2009**, *394*, 2119.
[16] Kieken F, Pinel G, Antignac JP, Garcia P, Popot MA, Grall M, *et al*. Development and application of a metabolomic approach based on liquid chromatography-high resolution mass spectrometry to reveal an illegal administration of recombinant equine growth hormone in horse from urinary and plasmatic biological signatures. *Metabolomics* **2011**, *7*, 84.
[17] F. Boyard-Kieken, G. Dervilly-Pinel, P. Garcia, A.-C. Paris, M.-A. Popot, B. Le Bizec, *et al*. Comparison of different liquid chromatography stationary phases in LC-HRMS metabolomics for the detection of recombinant growth hormone doping control. *J. Sep. Sci.* **2011**, *34*, 1.
[18] G. Dervilly-Pinel, S. Weigel, A. Lommen, S. Chereau, L. Rambaud, M. Essers, *et al*. Assessment of two complementary LC-HRMS metabolomics strategies for the screening of anabolic steroid treatment in calves. *Anal. Chim. Acta* **2011**, *700*, 144.
[19] J.K. Nicholson, J.C. Lindson. Systems biology. *Nature* **2008**, *455*, 1054.
[20] S.G. Oliver, M.K. Winson, D.B. Kell, F. Baganz. Systematic functional analysis of the yeast genome. *Trends Biotechnol.* **1998**, *16*, 373.
[21] J. Nicholson, J. Lindon, E. Holmes. Metabonomics: Understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotics* **1999**, *29*, 1181.
[22] J. Nicholson, J. Connelly, J. Lindon, E. Holmes. Metabonomics: A platform for studying drug toxicity and gene function. *Nat. Rev. Drug Discov.* **2002**, *1*, 153.
[23] D. Robertson. Metabonomics in toxicology: A review. *Toxicol. Sci.* **2005**, *85*, 809.
[24] D. Wishart. Metabolomics: Applications to food science and nutrition research. *Trends Food Sci. Technol.* **2008**, *19*, 482.
[25] J.-P. Antignac, F. Courant, G. Pinel, F. Monteau, B. Le Bizec. Mass spectrometric based metabolomic in the field of chemical risk: state of the art, pitfalls, and future trends. *Trends Anal. Chem.* **2011**, *30*, 292.
[26] X. Cai, L. Zou, J. Dong, L. Zhao, Y. Wang, Q. Xu, *et al*. Analysis of highly polar metabolites in human plasma by ultra-performance hydrophilic interaction liquid chromatography coupled with quadrupole time of flight mass spectrometry. *Anal. Chim. Acta* **2009**, *650*, 10.
[27] J.B. German, L.A. Gillies, J.T. Smilowitz, A.M. Zivkovic, S.M. Watkins. Lipidomics and lipid profiling in metabolomics. *Curr. Opin. Lipidol.* **2007**, *18*, 66.
[28] C. Legido-Quigley, C. Stella, F. Perez-Jimenez, J. Lopez-Miranda, J. Ordovas, J. Powell, *et al*. Liquid chromatography-mass spectrometry methods for urinary biomarker detection in metabonomic studies with application to nutritional studies. *Biomed. Chromatogr.* **2010**, *24*, 737.
[29] R. El Rammouz, F. Letisse, S. Durand, J.C. Portais, Z.W. Moussa, X. Fernandez. Analysis of skeletal muscle metabolome: evaluation of extraction methods for targeted metabolite quantification using liquid chromatography tandem mass spectrometry. *Anal. Biochem.* **2010**, *398*, 169.
[30] B. Alvarez-Sanchez, F. Priego-Capote, M. Luque de Castro. Metabolomics analysis II. Preparation of biological samples prior to detection. *Trends Anal. Chem.* **2010**, *29*, 120.
[31] J. Gullberg, P. Jonsson, A. Nordstrom, M. Sjostrom, T. Moritz, *Anal. Biochem.* **2004**, *331*, 283.
[32] B. Alvarez-Sanchez, F. Priego-Capote, M. Luque de Castro. Metabolomics analysis I: Selection of biological samples and practical aspects preceeding sample preparation. *Trends Anal. Chem.* **2010**, *29*, 111.
[33] M.-E Dumas, L. Debrauwer, L. Beyet, D. Lesage, F. André, A. Paris, *et al*. Analyzing the physiological signature of anabolic steroids in cattle urine using pyrolysis/metastable atom bombardment mass spectrometry and pattern recognition. *Anal. Chem.* **2002**, *74*, 5393.
[34] M.-E. Dumas, C. Canlet, J. Vercauteren, F. André, A. Paris. Homeostatic Signature of Anabolic Steroids in Cattle Using 1H-13C HMBC NMR Metabonomics. *J. Proteome Res.* **2005**, *4*, 1493.
[35] E.J. Jarjalainen. The spectrum reconstruction problem. Use of alternating regression for unexpected spectral components in two dimensional spectroscopies. *Chemometr. Intell. Lab. Syst.* **1989**, *7*, 31.
[36] M. Katajamaa, J. Miettinen, M. Oresic. MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. Bioinformatics Application Note 22. **2006**, pp. 634.
[37] C.A. Smith, E.J. Want, G.C. Tong, A. Saghatelian, B.F. Cravatt, R. Abagyan, *et al*. XCMS: Processing mass spectrometry data for metabolite profiling using non linear peak alignment, matching and identification. *Anal. Chem.* **2006**, *78*, 779.
[38] A. Lommen. MetAlign: Interface-driven, versatile metabolomics toll for hyphenated full scan mass spectrometry data processing. *Anal. Chem.* **2009**, *81*, 3079.
[39] C. Broeckling, I. Reddy, A. Duran, X. Zhao, L. Summer. MET-IDEA: data extraction tool for mass spectrometry-based metabolomics. *Anal. Chem.* **2006**, *78*, 4334.
[40] M.M.W. Hendriks, F.A. van Eeuwijk, R.H. Jellema, J.A. Westerhuis, T.H. Reijmers, H.C.J. Hoefsloot, *et al*. Data-processing strategies for metabolomics studies. *Trends Anal. Chem.* **2011**, DOI: 10.1016/j.trac.2011.04.019
[41] T. Kind, O. Fiehn. Seven golden rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics* **2007**, *8*, 105.
[42] World anti doping code. article 2.1.2, December **2009**.
[43] International Agreement on Breeding, Racing and Wagering, article 6, April **2009**.
[44] G. Liumbruno, A. D'Alessandro, G. Grazzini, L. Zolla. Blood related proteomics. *J. Proteomics* **2010**, *73*, 483.
[45] S. Rochereau-Roulet, I. Gaudin, S. Chéreau, S. Prévost, G. André-Fontaine, G. Pinel, *et al*. Development and validation of an enzyme-linked immunosorbent assay for the detection of circulating antibodies raised against growth hormone as a consequence of rbST treatment in bovine. *Anal. Chim. Acta* **2011**, *700*, 189.
[46] F. Kieken. Mise au point d'un nouvel outil de dépistage de l'administration de l'hormone de croissance équine recombinante par approche métabolomique, PhD Thesis, **2010**, Nantes (France) University.