

New surveillance concepts in food safety in meat producing animals: the advantage of high throughput 'omics' technologies — A review

Michael W. Pfaffl^{1,a,*} and Irmgard Riedmaier-Sprenzel^{1,2,a}

* **Corresponding Author:** Michael W. Pfaffl
Tel: +49-8161-713511, **Fax:** +49-8161-713539,
E-mail: michael.pfaffl@wzw.tum.de

¹ Animal Physiology and Immunology, TUM School of Life Sciences, Technical University of Munich Weihenstephan, Weihenstephaner Berg 3, 85354 Freising, Germany

² Eurofins Medigenomix Forensik GmbH, Anzinger Straße 7a, 85560 Ebersberg, Germany

^a These authors contributed equally to this work.

ORCID

Michael W. Pfaffl
<https://orcid.org/0000-0002-3192-1019>

Submitted Mar 7, 2018; Accepted May 23, 2018

Abstract: The misuse of anabolic hormones or illegal drugs is a ubiquitous problem in animal husbandry and in food safety. The ban on growth promotants in food producing animals in the European Union is well controlled. However, application regimens that are difficult to detect persist, including newly designed anabolic drugs and complex hormone cocktails. Therefore identification of molecular endogenous biomarkers which are based on the physiological response after the illicit treatment has become a focus of detection methods. The analysis of the 'transcriptome' has been shown to have promise to discover the misuse of anabolic drugs, by indirect detection of their pharmacological action in organs or selected tissues. Various studies have measured gene expression changes after illegal drug or hormone application. So-called transcriptomic biomarkers were quantified at the mRNA and/or microRNA level by reverse transcription-quantitative polymerase chain reaction (RT-qPCR) technology or by more modern 'omics' and high throughput technologies including RNA-sequencing (RNA-Seq). With the addition of advanced bioinformatical approaches such as hierarchical clustering analysis or dynamic principal components analysis, a valid 'biomarker signature' can be established to discriminate between treated and untreated individuals. It has been shown in numerous animal and cell culture studies, that identification of treated animals is possible via our transcriptional biomarker approach. The high throughput sequencing approach is also capable of discovering new biomarker candidates and, in combination with quantitative RT-qPCR, validation and confirmation of biomarkers has been possible. These results from animal production and food safety studies demonstrate that analysis of the transcriptome has high potential as a new screening method using transcriptional 'biomarker signatures' based on the physiological response triggered by illegal substances.

Keywords: Livestock; Growth Promotants; Anabolic Hormones; Biomarker Signatures; Transcriptome; Drug Abuse Detection

INTRODUCTION

Application of growth promoting agents like steroid hormones, β -Agonists or anabolic hormone cocktails are well known to increase muscle mass in meat producing farm animals. The main purpose is to reduce lipid mass and increase muscle tissue. These agents promote overall tissue growth, by increasing growth rates, improve the conversion of nitrogen in feed into muscle protein, and hence increase the amount of lean and valuable meat in the carcass.

The application of any growth promoting agents is strictly forbidden in livestock production within the European Union (EU) [1]. Selected toxicological reports in meat products state that the use of a single anabolic substance has no effect on the consumer, and hence the use of these substances is licensed in some countries like USA, Brazil, Mexico, or South Africa [2]. However a previous report by Swan et al [3] described a reduction of sperm quality of male offspring that was positively correlated with the mother's consumption of beef during

pregnancy. This research showed a direct connection between the residues from application of growth promotants and their activity in meat. Such metabolites can influence the development of male reproductive organs *in utero*, providing supporting evidence for the ban on growth promoting drugs in European animal husbandry [3].

Today, these growth promoting drugs are screened within an EU control program, using sensitive immunological-based test systems (e.g. ELISA, enzyme-linked immunosorbent assays) or chromatographic methods in combination with mass spectrometry [4-6]. Using these established methods, newly designed anabolic drugs cannot be identified until their chemical structure is known, or a specific and high affinity antibody is available for the immune tests. A further problem is the identification of a single drug after the application of a drug cocktail, which may represent a broad mixture of various different anabolic substances, xenobiotics or bioactive metabolites, each in a low concentration. These undefined cocktails can have, in sum, a comparable biological effect similar to a single drug administered at a higher level. However, due to their low concentrations they are barely detect, and difficult to quantify in a reproducible manner using conventional detection methods [7]. Therefore, development of new, sensitive screening and quantification methods is required to detect use of such substances, independent of their structure and/or concentration.

In designing xenobiotic anabolic drugs, the chemical structures of these new agents are changed slightly, but so that the desired physiological effect based on the cellular drug-to-receptor binding remains functional. Hence the effect of xenobiotics or drug cocktails should be close to the natural physiological receptor-ligand action, in the case of livestock, increased weight gain, greater muscle mass, and a reduction in fat tissue. The observable cellular reaction after the illicit administration of growth promoters is an increase and change in physiological signalling altering animal physiology. The detection of increased and altered cellular signalling is a promising approach to develop new screening methods to identify the misuse of anabolic agents [8]. The challenge is how to easily quantify these physiological reactions at the cellular and molecular level?

In molecular biology laboratories, various methods are available to monitor physiological changes at the genomic, transcriptomic, proteomic and/or metabolomic level [9]. Those relatively new and innovative methods are known as 'omic' technologies. They can be differentiated into targeted and untargeted methods. With targeted methods it is possible to quantify single, selected and known factors, such as the quantification of a genetic marker, or the expression of a single RNA sequence or a protein. Untargeted methods screen for a large amount of differentially expressed factors in parallel at the same time [10]. In general, untargeted methods are used to screen for unknown factors, and to discover new potential bio-

markers (biomarker candidates) which can be later be verified and confirmed using targeted methods.

Among the so called 'omic' technologies, the analysis of the transcriptome provides a very promising approach to find new and reliable biomarkers in the field of anabolic agents, because it is known that steroid hormones directly, and β -agonists indirectly, influence the transcription of specific genes [8,11]. Nowadays, transcriptomics has progressed beyond the quantification of gene expression using analysis of mRNA. The analysis of regulatory small RNA species, namely the microRNAs, is becoming increasingly important, especially to identify disease specific biomarkers [12]. MicroRNAs are naturally expressed in all tissues and organs, and are involved in the post-transcriptional regulation of mRNA gene expression by inhibiting gene expression or translation, or by blocking or neutralizing targeted mRNA molecules. Hence, they play a major regulatory role in any cellular biological process, and show an individual and distinct expression pattern under different physiological conditions, disease states, or after the administration of drugs. Due to these properties, the analysis of microRNA expression changes is an interesting and promising approach to find a valid 'biomarker signature' to detect anabolic misuse [8,12].

This review describes new surveillance concepts for food safety in meat producing animals, by explaining the technical and analytical requirements of mRNA and microRNA gene expression profiling. Further, the advantages of the high throughput 'omics' technologies are described, as are efforts being made to identify transcriptional biomarkers to detect misuse of anabolic agents in cattle.

ANALYTICAL METHODS TO DETECT AND QUANTIFY GENE EXPRESSION CHANGES – FROM THE BASICS TO HIGH THROUGHPUT METHODS

The transcriptome is the complete set of RNA transcripts present in a particular cell, and the most prominent candidates investigated in research are the messenger RNAs (mRNA), small-RNAs (in particular the micro RNAs), transfer RNAs (tRNA), and ribosomal RNAs (rRNA). Transcriptomics describes the global study of gene expression at a certain point in time, for example as a reaction after a specific treatment. The basic methods available for gene expression profiling and to quantify gene expression can be subdivided into two groups: untargeted and targeted analytical methods [7,11]. The untargeted methods use 'hypothesis free' analysis of gene expression changes by analysing all expressed genes. This means that all changes at the transcriptomic level that are caused by anabolic treatment are detected within one experiment. Due to the large amount of gene expression data produced, the statistical analysis of these experiments is very time consum-

ing, requires specialized software packages, and a high degree of knowledge. With gene expression microarrays, a very large number of genes spotted to the microarray can be analysed, either together in a comparative way or separately, side-by-side. Therefore DNA fragments of all identified and known genes (called probes which represent specific coding regions of target genes) are immobilized on the solid surface (e.g. glass slide, synthetic membrane) of the microarray, whereby one entire gene is represented by more than one probe sequence and one probe is present in multiple copies [11,13]. Analyses of the gene expression changes in a comparative way between two biological samples, requires both total RNA pools to be reverse transcribed to cDNA and thereby labelled with two different coloured fluorescent dyes. Later, the two samples are mixed then hybridized to the slide whereupon the cDNAs bind to their complementary probes. The amount of bound cDNA copies can be quantified by the intensity of emission of the differently labelled fluorescent dyes [13]. Using specific fluorescent scanners and data analysis software, the gene expression data can be analysed and the changes and differences between the two samples can be quantified. Such comparative microarrays have until now been the 'gold standard' method to screen all gene expression differences between two biological samples. Currently, there are also microarrays available for the identification of small-RNA families, such as microRNAs. A variety of customized mRNA or microRNA microarrays are also available with selected target genes or microRNAs of interest for specific biological questions or physiological pathways of high interest (e.g. immune system, cardiovascular diseases, or cancer related) [14]. A major disadvantage of the microarray technology is that only the expression of genes with sequences already known and present in the gene databases can be analysed. A further problem is high basal background level, which reduces the signal-to-noise ratio and limits the range of detection, especially if one wants to detect lowly expressed genes of interest [15].

Serial analysis of gene expression (SAGE) is another example of an untargeted gene expression method. This technique is based on the generation of unique sequence tags of about 9 to 10 base pairs (bp) in length from mRNAs present, and the joining of these tags to long molecules for sequencing [16]. A length of 9 to 10 bp is sufficient to distinguish up to 4^9 (262,144) to 4^{10} (1,048,576) transcripts. Today, SAGE is an outdated sequencing technology, because it is very time consuming due to the number of working steps, the ligation of multiple short sequence tags, and later cloning into a vector to enable sequencing of numerous mRNAs within one and the same reaction. This method uses oligo dT primers for reverse transcription (RT) and hence can only be used for the identification of long RNAs such as mRNAs [16].

Next generation sequencing (NGS) is the current sequencing method of choice, representing a fast and reliable technology

for deep sequencing of any nucleic acid. With this innovative technology whole genomes and transcriptomes can be rapidly sequenced. In the following section we will focus on RNA Sequencing, the holistic analysis of small and/or long RNA families. From the total RNA sample, a cDNA library is created by attaching a subset of specific adapters to every RNA fragment. Each of these generated fragments is sequenced in a high throughput manner, resulting in a high number of short sequence reads that will either be aligned to a reference genome or assembled *de novo* without knowing the genome sequence of the species [15,17,18]. With this approach small-RNA fragments such as microRNA or piRNAs can also be analysed [15,19]. RNA sequencing has no sequencing limitation or upper limit of quantification breadth, and allows the identification of new sequences of any organism being studied. It has a high dynamic range of expression levels from single to millions of reads, has barely any background signal, and allows for quantification of multiple biological samples within a multiplexing set-up [15,20].

For the analysis of a fixed set of genes the targeted methods are referred. The first method developed in this field was the 'northern blot' analysis. Northern blotting involves the electrophoretic length separation of isolated RNA and the capillary transfer of RNA from the agarose gel to a positive charged blotting membrane. Identification and visualisation of individual target RNAs is enabled by RNA or DNA hybridisation probes composed of labelled nucleic acids, with a sequence complementary to the target RNA. Such hybridisation probes are labelled with different detection markers with either radioactive isotopes (mainly P^{32}), fluorescence dyes or chemiluminescence [21].

Today the 'gold standard' in single target gene expression analysis is reverse transcription-quantitative polymerase chain reaction (RT-qPCR). This method is based on RT followed by a pPCR. RT is used to reverse transcribe RNA into complementary DNA (cDNA) which is the starting template of a PCR. PCR is then applied to amplify defined parts of the DNA bordered by the two primers using DNA polymerases in a defined 3-step temperature protocol. Today, real time RT-qPCR enables the amplification and subsequent monitoring of the amplified DNA product during the course of the reaction [22,23].

Nowadays, multiplex RT-qPCR setup enables the quantification of a 'set of genes' in parallel and in one reaction vessel. Therefore, DNA probes for the different genes of interest are labelled with different coloured dyes emitting light measurable at different wave lengths [24,25]. Another high throughput application is customized RT-qPCR panels provided in a 96- or 384 well-plate format. These panels combine either a set of target genes present in one particular biological pathway, or allow the quantification of multiple transcription biomarkers to produce a so-called biomarker signature [11,26]. These

panels have fixed reaction components in all wells and are designed to run at common RT and PCR conditions. Hence, many target genes from one biological sample can be quantified in parallel within one run, enabling screening for gene expression changes of multiple genes. Such RT-qPCR setups are also available for the quantification of specific microRNA panels [26,27].

HOW TO USE THE TRANSCRIPTOME ANALYSIS FOR SURVEILLANCE CONCEPTS

To develop a reliable 'biomarker signature' based on a valid set of transcriptional biomarkers, various animal studies that produce reproducible 'real life' physiological effects are required [11,28]. Each drug or anabolic agent induces individual kinetic and tissue or organ transcription profile specific responses, which is based on different individual physiological and biological functions. 'Cell culture models' can be used as an alternative, but they are simplified models and only useful to study the physiological action of drugs in a tissue-specific context. Cell culture models consist mainly of one sub-group of cells and do not include interactions between different cell types, tissues or organs that occur in living organisms within animal studies.

Various studies exist that describe gene expression changes in a variety of organs caused by administration of different drug or anabolic substances in order to identify potential biomarkers. The choice of the most appropriate tissue or organ for screening and development of a 'biomarker signature' is essential. Each organ exhibits specific feedback and reaction kinetics after drug treatment, hence the physiological outcomes will be different between organs. The tissue should be easily accessible for sampling or, even better, non-invasive sampling using liquid biopsy is desirable [11,20]. With the focus on the administration of steroidal active drugs, the organs of the reproductive tract are the major focus due to their dependency on steroid hormones. In males, the testes and prostate are primary candidate tissues to analyse the influence of steroids on gene expression [11,28].

A LITERATURE SURVEY ON TRANSCRIPTIONAL BIOMARKER DISCOVERY

The potential power of transcriptional gene expression biomarkers for diagnostic use was first described for cancer detection [29]. Anabolic hormones exhibit a range of physiological effects, hence the search for valid gene expression biomarkers is a promising approach to develop screening methods for the use or misuse of anabolic hormones. There are scientific studies that show that steroid hormones and β -agonists

significantly affect gene expression in various organs. Reiter and coworkers [30] quantified changes in mRNA expression for a number of genes in *bovine* liver, muscle and uterus that are controlled by different xenobiotic anabolic agents. They found several regulated genes that were initial candidates for development of gene expression biomarkers [30]. The gene expression changes in heifers were caused by treatment with commercial growth promoters: melengestrol acetate (MGA), Finaplix-H (trenbolone acetate), or Ralgro (zeranol). Candidate genes chosen for their physiological action in different biological pathways were quantified via RT-qPCR. The study identified 8 to 18 significantly regulated genes in the tissues studied, although none of these genes was up- or down-regulated more than 2-fold. In the uterus, 13 significantly regulated genes were identified, and the expression changes were up to 6-fold [30]. These results indicate that steroid hormone dependent tissues are premier targets for measurement of gene expression and are optimally suited for biomarker discovery to screen for the misuse of anabolic steroids. In support of our hypothesis, the impact of steroids on the mRNA expression of several genes was shown by a number of research groups globally. Lopparelli et al [31] examined the expression profiles of 12 candidate genes via RT-qPCR in cattle testis after oral or intramuscular administration of dexamethasone (DEX), either alone or in combination with 17β -estradiol [31]. When DEX was applied intramuscularly, two genes were identified as being significantly regulated ($P450_{C17}$ and MR-like). Further, various members of the HSD17 family which regulate corticosteroid action, were regulated in animals that received the growth promoting drugs. In conclusion, no single biomarker was identified after the drug administration, but the first potential candidates were identified. In the same study, the influence of the prohormones dehydro-epiandrosterone (DHEA) and boldione on mRNA gene expression in testis was monitored. When DHEA was administered alone, it up-regulated the expression of aromatase and, in combination with boldione, DHEA significantly altered the expression of $P450_{SCC}$, 17β -hydroxysteroid dehydrogenase 1 and the androgen receptor [31]. In a further animal study using a combination of trenbolone acetate plus estradiol, gene expression changes were assessed in the reproductive tract of heifers. In the uterine endometrium and ovary, following treatment, 9 genes were identified as being significantly regulated in uterine horn and 12 genes in uterine corpus, whereas the expression of a further 4 genes (androgen receptor [AR], B-cell lymphoma 2 modifying factor 4, Caspase 3, and complement factor C7) were significantly regulated [32]. Comparing these results with those of Reiter et al [30,31], it can be seen that AR and Caspase 3 were also significantly regulated by Finaplix, indicating these genes could be suitable as the first biomarkers in uterine tissue to detect treatment with trenbolone acetate [30, 31]. Regarding the results obtained in the ovary within which

22 genes were significantly regulated, it can be concluded that ovary is an organ with more potential for the identification of gene expression changes in order to find biomarkers for the misuse of anabolic steroids [32]. Another new approach was the quantification of gene expression in vaginal swabs containing vaginal epithelial cells. This swab can be taken in a noninvasive form from a living individual. The vaginal epithelium has proven to be an optimal target tissue to identify direct hormone responsiveness to estrogens. Hence the changes in expression of numerous genes, especially of factors involved in keratinization, cell growth and apoptosis could be expected. From the set of 27 analyzed genes, 13 genes were identified as significantly regulated, and nearly all genes studied showed a trend to be regulated but did not reach the level of significance [33].

A further target tissue of focus is skeletal muscle. It is relatively easy to access by biopsy and is under direct control of growth promoting agents. De Jager et al [34] used hybridization microarrays to identify differently expressed genes in *longissimus dorsi* muscle after implantation of Revalor-H. In this experiment, 121 out of 16,944 quantified genes were differentially expressed. Surprisingly, one gene that previously not expected to be involved in anabolic functions that was shown to be highly regulated in this study was oxytocin. It was highly, significant upregulated 97-fold compared to the expression in control cattle. The plasma levels of oxytocin protein were also measured and found to be increased by 50-fold, making oxytocin a new and very promising biomarker candidate at the mRNA and protein levels [34]. Toffolatti and coworkers [35] treated veal calves either with a combination of estradiol benzoate plus testosterone enantate or estradiol benzoate plus boldenone undecylenate, and discovered gene expression changes by using RT-qPCR. 3-Hydroxy-3-methylglutaryl-CoA synthase 1 and AR were regulated in testis of all treated animals and on treatment days, and therefore these two genes represent potential biomarker candidates in testis to detect administration of an estrogen and androgen cocktail [35].

Further promising candidate genes for the development of a screening method in muscle and liver of cattle is the family of steroid hormone receptors in multiple tissues (e.g. liver, muscle, uterus, the gastrointestinal tract, kidney, vagina, prostate and blood cells), Insulin-like growth factor-1 (IGF-1), and various pro- and anti-inflammatory, apoptotic and proliferative genes in blood cells [36]. Further, growth promoting β -agonists are known to affect the expression profile of muscle proteins including α -actin, myosin and calpastatin in cattle. The mRNA expression profile of β -adrenergic receptors are also known to be influenced by their natural or xenobiotic ligands [37].

A further tissue of key interest is the liver. It is involved in various hormonally-regulated pathways and in drug detoxification [38,39]. Hence it is an attractive tissue for the identification of potential biomarkers to screen for the use

of anabolic agents. Giantin et al [40] analyzed the influence on gene expression of DEX administered orally or intramuscularly, either alone or in combination with 17β -estradiol, in *bovine* liver. They identified 4 significantly regulated genes (cytochrome P450 family 2 subfamily B member 6, cytochrome P450 family 2 subfamily E member 1 [CYP2E1], glutathione S-transferase alpha 1, and sulfotransferase family 1A member 1) as valuable biomarker candidates to identify the use of DEX and estradiol [40]. Rijk and coworkers [41] performed a microarray experiment on the influence of DEX administered orally or intramuscularly on gene expression in liver. Twenty-three genes were identified whose expression was altered by the oral administration of DEX, and expression of 37 genes was altered when administration was performed intramuscularly. Only one of these genes, deleted in malignant brain tumors-1 was regulated in both groups [41]. Becker et al [38] analyzed expression changes in the liver of heifers treated with trenbolone acetate plus estradiol. They quantified the expression of 34 genes belonging to metabolic pathways controlled by steroid hormone action, of which 11 genes were significantly regulated [38]. They also tried to quantify changes in microRNA expression using a ready to use PCR panel, enabling the screening of 730 microRNAs. It turned out that 36 microRNAs showed significant regulation after drug application [27], demonstrating potential of the additional analysis of microRNAs to identify new biomarkers for the misuse of anabolic substances. In a further animal study, where calves were treated with a combination of estradiol and progesterone or clenbuterol (CLEN), respectively, Riedmaier et al [39] identified using a principal components analysis (PCA), a signature comprising 11 genes (adrenoceptor beta 2, adenylate kinase 3 L1, apolipoprotein A4 [APOA4], apolipoprotein C2, estrogen receptor 2, F-box protein 39, glycerophosphodiester phosphodiesterase domain containing 1, high mobility group Coenzyme A reductase, insulin receptor 1, paraoxonase 1, and tyrosine aminotransferase) that enabled the separation of treated from untreated animals independent of the applied drug [39]. They assessed the untreated control animals across a broader range of physiological conditions and showed that their biomarker signature successfully separated treated from untreated animals.

The studies described above from various research groups around the world clearly show that the development of a valid 'biomarker signature' at the mRNA or microRNA transcriptome level, based on physiological effects and independent of the administered substance is possible.

NONINVASIVE AND MINIMAL INVASIVE SAMPLING

The fact that sampling of most of the tissues and organs described above can only be taken after slaughter is a big disadvantage

for screening. Hence these tissues do not present as promising primary target tissues for developing a screening method for illicit drug administration. To enable controls on-farm, tissues or matrices that can be sampled from living individuals in a noninvasive form and independent of gender must be the focus of further research. Reiter et al [42] monitored gene expression changes in cell culture experiments with human hair follicle dermal papilla cells that were treated with stanozolol [42]. In a second *in vivo* study, it was shown that it is possible to extract RNA from hair follicle cells and that gene expression in these cells is also influenced by steroid hormones [43]. In conclusion, hair roots represent a promising tissue to find gene expression biomarkers with potential to develop a noninvasive screening method based on gene expression patterns.

Second, blood is also a potential minimal-invasive tissue matrix for sampling. There are various published studies describing the effects of growth promoting agents on gene expression in blood or in blood-derived white blood cells [44]. However, most of these experiments were performed *in vitro* in cell culture. Cantiello et al [7] described the influence of administering a hormone cocktail of 17 β -estradiol, DEX and CLEN on the expression of 4 different cytokines in primary blood lymphocytes, with only interferon gamma significantly regulated [7]. The effects of trenbolone acetate in combination with estradiol on the expression of 36 candidate genes across different time points post-administration was quantified by Riedmaier et al [44]. Eleven of the candidate genes were identified as significantly regulated, although only two genes, glucocorticoid receptor α and interleukin 1 alpha, were regulated at two time points [44].

WITH THE HELP OF BIOINFORMATICS

The above studies show that the identification of a single, dominant gene expression biomarker for the misuse of different anabolic agents is difficult. A more promising approach is the identification of a biomarker pattern, a combination of a set of regulated genes summarized as a 'biomarker signature' [19].

In biomarker discovery, a pattern of biomarkers emerges, with multiple factors being influenced quantitatively by a drug or specific physiological stage. The most important question is how to deal with a huge data set in order to extract, interpret and visualize useful information. Regardless of which *omic* technology is used for biomarker research, bioinformatical tools are necessary to deal with the data and extract the needed essential information. To transform the high-dimensional data into a reduced subspace that represents data in far fewer dimensions, methods for dimensionality reduction are used [8,45]. In combination with pattern recognition technologies, the identification and visualization of the desired information is necessary.

A simple method to classify samples using transcriptome expression patterns is two- or three-dimensional scatter plots. Using this method, only two or three transcripts, proteins or metabolites are assessed. If more components need to be taken into account, multivariate analytical methods are required. The PCA reduces multidimensional data sets to fewer dimensions called 'principal components' (PC) [46]. Each sample analyzed is visualized by one spot that results from diminishing all collected data for a specific sample into two principle components, and so that each sample analyzed is represented by one spot. Employing this method for growth promotant treatment screening will, ideally, result in a graph with two groups of spots; one group represents the untreated controls, and the other group the treated individuals. PCA was effectively used by Becker et al [32] and Riedmaier et al [39] to identify potential gene expression biomarker patterns for anabolic treatment using *bovine* blood, vaginal epithelial cells and liver [32,39]. A further method for visualizing treatment patterns based on multivariate data is hierarchical cluster analysis (HCA). The hierarchical order is represented by a tree dendrogram in which related samples are more closely clustered together than samples that are more different [46]. Used in anabolic treatment screening, HCA should result in a tree in which the treated or the untreated samples, respectively, are close together, and the group of treated samples is separated from the group of untreated samples. Regardless of which biostatistical method is employed for treatment screening, it is always necessary to have a large number of untreated controls serving as the basis for physiologically normal individuals. High biological variance between each animal are caused by genetic or environmental diversity. To deal with these differences between individuals, a large number of control samples representative of the genetic and environmental variation within the investigated group of animals are therefore needed.

In summary, the extraction of the useful information from a data set using biostatistical methods for pattern recognition such as PCA or HCA is particularly helpful [46]. Such methods have already proven successful using *bovine* uterus, ovary, vaginal smear, liver and blood to visualize separation between untreated control and treated animals [8,11,27,38,39]. These results show that the complex analysis of gene expression changes using biostatistical pattern recognition methods is a promising approach to identify biomarker patterns to screen for the misuse of anabolic agents in cattle. In recent years much effort has been made in the analysis of gene expression patterns to discover and identify 'biomarker signatures' for the screening of illicit growth promotant use in livestock.

THE OPTIMAL COMBINATION – RNA-Seq AND BIOINFORMATICS

The use of '-omic' technologies is a highly discussed, new ap-

proach for the identification of the abuse of anabolic agents in sport and also in farm animals [11,20,28]. With these up-to-date technologies, physiological changes caused by the application of illicit substances can be detected at the molecular level. The most promising method is the detection of gene expression changes, because the key substances used for anabolic purposes - steroid hormones and β -agonists - are known to directly influence the expression of specific growth-relevant genes. Since, the recent development of high throughput sequencing technologies has enabled the establishment of a holistic gene expression profiling, called RNA-sequencing (RNA-Seq). This method allows the detection of all expressed mRNAs, or other long non-coding RNAs, and is able to detect 'one single RNA molecule' and is thereby nearly as sensitive as RT-qPCR. Compared with the microarray method, RNA-Seq has no upper limit of detection, shows a higher dynamic range of expression levels and has nearly no background signal [15,20]. A further fundamental advantage of this new approach is the possibility of *de novo* RNA detection, enabling the identification of new mRNAs, RNA families or new splice

variants of expressed genes [47]. Riedmaier and coworkers applied RNA-Seq to very successfully screen for potential gene expression biomarkers and to detect illegal use of anabolic agents in farm animals [28]. In heifers, expression analysis of 40 candidate genes chosen by RNA-Seq resulted in dynamic, significantly regulated genes with extreme fold regulation values, ranging from 5-fold downregulation (APOA4) up to 82-fold up-regulation (SerpinI2). This shows that hypothesis free gene expression screening by RNA-Seq results in more pronounced and more highly regulated genes than by choosing candidates using a literature research. This technology enables quantification of changes in the expression of unexpected genes or new splice variants, and the influence of treatment with anabolic agents on pathways that are not yet known to be influenced can also be detected. To also verify these candidate markers in other species, their expression was quantified in liver samples from boars treated with Synovex plus (trenbolone acetate plus estradiol). It was shown that 4 genes were significantly regulated, with expression dynamics of a 25-fold downregulation of CYP2E1 and a 5-fold upreg-

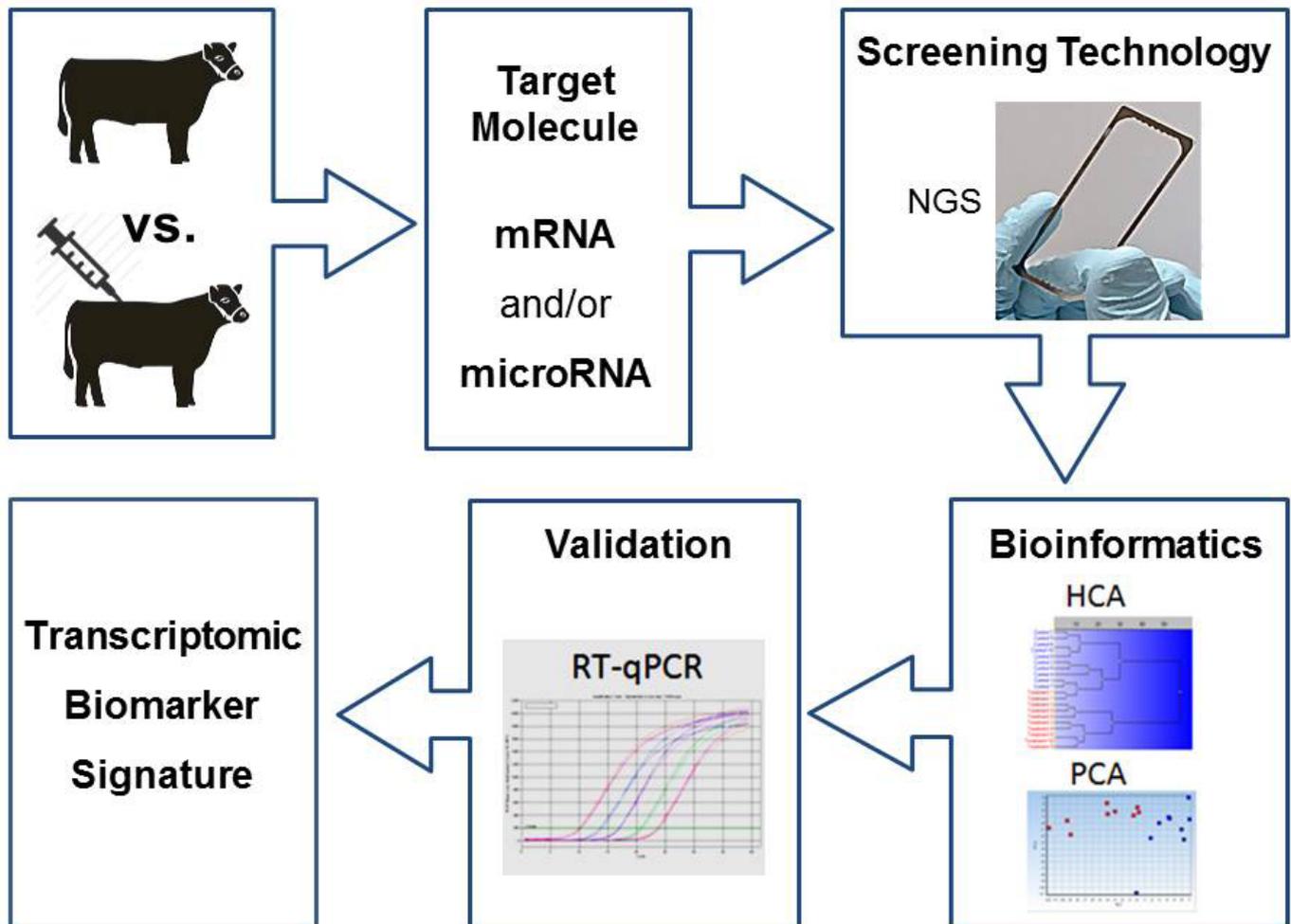


Figure 1. Workflow for the detection of growth promoting and anabolic substances in beef cattle based on the screening, identification, and validation of a transcriptomic biomarker signature.

ulation of cut like homeobox 2. The biostatistical analysis via PCA and HCA resulted in a complete separation of treated and untreated animals [28].

The results obtained in these heifer and boar studies indicate that identified regulated genes represent initial biomarker candidates for the detection of treatment with a combination of trenbolone acetate plus estradiol in farm animals, independent of species, breed, sex, reproductive state or age.

CONCLUSION

The determination of physiological changes caused by the use of growth promoting and anabolic substances based on the analysis of transcriptomic gene expression changes is a very promising method for the development of new and sensitive screening methods to combat against the misuse of anabolic agents in cattle.

This review demonstrates that RNA-Seq is the most up-to-date untargeted method to screen for regulated genes that may be used as biomarker candidate genes (Figure 1). Analyses of the resulting 'big' transcriptomics data set using biostatistical methods for pattern recognition (e.g. HCA or PCA) helps to extract the required information and to visualize separation between untreated controls and treated animals. An additional validation of the identified candidates via a targeted assay is recommended, classically done using RT-qPCR. The analysis of microRNA in a variety of tissues has also been shown to be a promising alternative. In particular, circulating microRNA samples using liquid blood biopsy have already been successfully shown to identify potential biomarkers.

For the future, more animal studies in which different growth promoters and drug cocktails are administered are necessary in order to validate existing transcriptional biomarkers and to further discovery of new candidates. In addition, more data are needed to define a normal basal expression profile from healthy, non-treated animals across a broad range of factors including breed, age, immune status, housing conditions, etc. The evaluation of gene expression changes in other matrices, tissues or biological fluids sampled in a noninvasive way from the living animal, such as blood, urine, sweat, or swabs, are of future importance for illicit drug use detection.

The fact that biomarkers identified in cattle and calves could be verified in pigs has led to the hypothesis that gene expression biomarkers are more or less 'independent' of the species, and therefore genes whose expression is altered in animals can also act as potential biomarkers for the detection of the misuse of anabolic agents in human sports.

CONFLICT OF INTEREST

We certify that there is no conflict of interest with any financial organization regarding the material discussed in the manu-

script.

ACKNOWLEDGMENTS

This article is dedicated to Prof. Heinrich H. D. Meyer (deceased April 13, 2012) whose passion was the field of residue analysis and who always believed in the idea of gene expression biomarkers to combat the abuse of anabolic substances. A special thanks to the Kontrollgemeinschaft Deutsches Kalbfleisch, the Société Générale de Surveillance (SGS) and Denkavit for supporting our animal study. Thank you also to the Genomics Core Facility, EMBL, Heidelberg Germany, especially Vladimir Benes for RNA-Sequencing analysis of our first samples. Thank you to the RIKILT Institute of Food Safety, Wageningen, The Netherlands for supporting various collaborative studies in calves and heifers.

REFERENCES

1. Council Directive 96/23/EC (OJ L125, p10, 23/05/1996) of 29 April 1996 on measures to monitor certain substances and residues in live animals and animal products. Brussels, Belgium: European Commission, 1996.
2. De Brabander HF, Le Bizec B, Pinel G, et al. Past, present and future of mass spectrometry in the analysis of residues of banned substances in meat-producing animals. *J Mass Spectrom* 2007;42:983-98.
3. Swan SH, Liu F, Overstreet JW, Brazil C, Skakkebaek NE. Semen quality of fertile US males in relation to their mothers' beef consumption during pregnancy. *Hum Reprod* 2007;22: 1497-502.
4. Meyer HHD, Hoffmann S. Development of a sensitive micro-titration plate enzyme-immunoassay for the anabolic steroid trenbolone. *Food Addit Contam* 1987;4:149-60.
5. Meyer HHD, Rinke L, Duersch I. Residue screening for the beta-agonists clenbuterol, salbutamol and cimaterol in urine using enzyme immunoassay and high-performance liquid chromatography. *J Chromatogr B Biomed Sci Appl* 1991;564: 551-6.
6. Scippo ML, Degand G, Duyckaerts A, Maghuin-Rogister G, Delahaut P. Control of the illegal administration of natural steroid hormones in the plasma of bulls and heifers. *Analyst* 1994;119:2639-44.
7. Cantiello M, Carletti M, Cannizzo FT, et al. Effects of an illicit cocktail on serum immunoglobulins, lymphocyte proliferation and cytokine gene expression in the veal calf. *Toxicology* 2007; 242:39-51.
8. Riedmaier I, Pfaffl MW. Transcriptional Biomarkers - High throughput screening, quantitative verification and bioinformatical validation methods. *Methods* 2013;59:3-9.
9. Pfaffl MW. Advanced molecular diagnostics for biomarker discovery. *Biomol Detect Quantif* 2015;5:1-2.

10. Pinel G, Weigel S, Antignac JP, et al. Targeted and untargeted profiling of biological fluids to screen for anabolic practices in cattle. *Trends Analyt Chem* 2010;29:1269-80.
11. Riedmaier I, Becker C, Pfaffl MW, Meyer HHD. The use of *omic* technologies for biomarker development to trace functions of anabolic agents. *J Chromatogr A*. 2009;1216:8192-9.
12. Heneghan HM, Miller N, Kerin MJ. MiRNAs as biomarkers and therapeutic targets in cancer. *Curr Opin Pharmacol* 2010;10:543-50.
13. Kurian KM, Watson CJ, Wyllie AH. DNA chip technology. *J Pathol* 1999;187:267-71.
14. Weile J, Knabbe C. Current applications and future trends of molecular diagnostics in clinical bacteriology. *Anal Bioanal Chem* 2009;394:731-42.
15. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 2009;10:57-63.
16. Yamamoto M, Wakatsuki T, Hada A, Ryo A. Use of serial analysis of gene expression (SAGE) technology. *J Immunol Methods* 2001;250:45-66.
17. Spornraft M, Haase B, Benes V, Pfaffl MW, Riedmaier I. Optimization of extraction of circulating RNAs from plasma – enabling small RNA sequencing. *PLoS One* 2014;9:e107259.
18. Costa V, Angelini C, De F, Ciccociolla A. Uncovering the complexity of transcriptomes with RNA-Seq. *J Biomed Biotechnol* 2010;2010:Article ID 853916.
19. Spornraft M, Kirchner B, Pfaffl MW, Riedmaier I. The potential of circulating extracellular small RNAs (smexRNA) in veterinary diagnostics – identifying biomarker signatures by multivariate data analysis. *Biomol Detect Quantif* 2015;5:15-22.
20. Buschmann D, Haberberger A, Kirchner B, et al. Towards reliable biomarker signatures in the age of liquid biopsies – how to standardize the small RNA-Seq workflow. *Nucleic Acids Res* 2016;44:5995-6018.
21. Alwine JC, Kemp DJ, Stark GR. Method for detection of specific RNAs in agarose gels by transfer to diazobenzylxymethyl-paper and hybridization with DNA probes. *Proc Natl Acad Sci USA* 1977;74:5350-4.
22. Higuchi R, Dollinger G, Walsh PS, Griffith R. Simultaneous amplification and detection of specific DNA-sequences. *Biotechnology (NY)* 1992;10:413-7.
23. Kubista M, Andrade JM, Bengtsson M, et al. The real-time polymerase chain reaction. *Mol Aspects Med* 2006;27:95-125.
24. Smith CJ, Osborn AM. Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiol Ecol* 2009;67:6-20.
25. Gunson RN, Bennett S, Maclean A, Carman WF. Using multiplex real time PCR in order to streamline a routine diagnostic service. *J Clin Virol* 2008;43:372-5.
26. Meyer SU, Kaiser S, Wagner C, Thirion C, Pfaffl MW. Profound effect of profiling platform and normalization strategy on detection of differentially expressed microRNAs – a comparative study. *PLoS ONE* 2012;7:e38946.
27. Becker C, Riedmaier I, Reiter M, et al. Changes in the miRNA profile under the influence of anabolic steroids in bovine liver. *Analyst* 2011;136:1204-9.
28. Riedmaier I, Benes V, Blake J, et al. RNA-sequencing as useful screening tool in the combat against the misuse of anabolic agents in cattle. *Anal Chem* 2012;84:6863-8.
29. Ginsburg GS, Haga SB. Translating genomic biomarkers into clinically useful diagnostics. *Expert Rev Mol Diagn* 2006;6:179-91.
30. Reiter M, Walf VM, Christians A, Pfaffl MW, Meyer HHD. Modification of mRNA expression after treatment with anabolic agents and the usefulness for gene expression-biomarkers. *Anal Chim Acta* 2007;586:73-81.
31. Lopparelli RM, Zancanella V, Giantin M, et al. Steroidogenic enzyme gene expression profiles in the testis of cattle treated with illicit growth promoters. *Steroids* 2011;76:508-16.
32. Becker C, Riedmaier I, Reiter M, et al. Influence of anabolic combinations of an androgen plus an estrogen on biochemical pathways in bovine uterine endometrium and ovary. *J Steroid Biochem Mol Biol* 2010;125:192-201.
33. Riedmaier I, Reiter M, Tichopad A, Pfaffl MW, Meyer HHD. The potential of bovine vaginal smear for biomarker development to trace the misuse of anabolic agents. *Exp Clin Endocrinol Diabetes* 2011;119:86-94.
34. De Jager N, Hudson NJ, Reverter A, et al. Chronic exposure to anabolic steroids induces the muscle expression of oxytocin and a more than fiftyfold increase in circulating oxytocin in cattle. *Physiol Genomics* 2011;43:467-78.
35. Toffolatti L, Rosa GL, Patarnello T, et al. Expression analysis of androgen-responsive genes in the prostate of veal calves treated with anabolic hormones. *Domest Anim Endocrinol* 2006;30:38-55.
36. Pfaffl MW, Daxenberger A, Hageleit M, Meyer HHD. Effects of synthetic progestagens on the mRNA expression of androgen receptor, progesterone receptor, estrogen receptor alpha and beta, insulin-like growth factor (IGF)-1 and IGF-1 receptor in Heifer tissues. *J Vet Med A Physiol Pathol Clin Med* 2002;49:57-64.
37. Sato S, Nomura S, Kawano F, et al. Effects of the beta2-agonist clenbuterol on beta1- and beta2-adrenoceptor mRNA expressions of rat skeletal and left ventricle muscles. *J Pharmacol Sci* 2008;107:393-400.
38. Becker C, Riedmaier I, Reiter M, et al. Effect of trenbolone acetate plus estradiol on transcriptional regulation of metabolism pathways in bovine liver. *Horm Mol Biol Clin Investig* 2011;2:257-65.
39. Riedmaier I, Spornraft M, Pfaffl MW. Identification of a potential gene expression biomarker signature in bovine liver to detect the abuse of growth promoters. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess* 2014;4:641-9.
40. Giantin M, Lopparelli RM, Zancanella V, et al. Effects of illicit dexamethasone upon hepatic drug metabolizing enzymes

- and related transcription factors mRNAs and their potential use as biomarkers in cattle. *J Agric Food Chem* 2010;58:1342-9.
41. Rijk JC, Peijnenburg AA, Hendriksen PJ, et al. Feasibility of a liver transcriptomics approach to assess bovine treatment with the prohormone dehydroepiandrosterone (DHEA). *BMC Vet Res* 2010;6:44.
 42. Reiter M, Pfaffl MW, Schönfelder M, Meyer HHD. Gene expression in hair follicle dermal papilla cells after treatment with stanozolol. *Biomark Insights* 2009;4:1-8.
 43. Reiter M, Lüderwald S, Pfaffl MW, Meyer HHD. First steps towards new screening method anabolic androgenic androgens in human hair follicles. *Doping J* 2008;5:3.
 44. Riedmaier I, Tichopad A, Reiter M, Pfaffl MW, Meyer HHD. Identification of potential gene expression biomarkers for the surveillance of anabolic agents in bovine blood cells. *Anal Chim Acta* 2009;638:106-13.
 45. Lee G, Rodriguez C, Madabhushi A. Investigating the efficacy of nonlinear dimensionality reduction schemes in classifying gene and protein expression studies. *IEEE/ACM Trans Comput Biol Bioinform* 2008;5:368-84.
 46. Bergkvist A, Rusnakova V, Sindelka R, et al. Gene expression profiling--Clusters of possibilities. *Methods* 2010;50:323-35.
 47. Costa V, Angelini C, De Feis I, Ciccodicola A. Uncovering the complexity of transcriptomes with RNA-Seq. *J Biomed Biotechnol* 2010;2010:Article ID853916.