Review

Drug Testing and Analysis

Received: 15 April 2011 Revised: 24 June 2011 Accepted: 25 June 2011 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI 10.1002/dta.331

Principle considerations for the use of transcriptomics in doping research

Elmo W. I. Neuberger, Dirk A. Moser and Perikles Simon*

ABSTRACT: Over the course of the past decade, technical progress has enabled scientists to investigate genome-wide RNA expression using microarray platforms. This transcriptomic approach represents a promising tool for the discovery of basic gene expression patterns and for identification of cellular signalling pathways under various conditions. Since doping substances have been shown to influence mRNA expression, it has been suggested that these changes can be detected by screening the blood transcriptome. In this review, we critically discuss the potential but also the pitfalls of this application as a tool in doping research. Transcriptomic approaches were considered to potentially provide researchers with a unique gene expression signature or with a specific biomarker for various physiological and pathophysiological conditions. Since transcriptomic approaches are considerably prone to biological and technical confounding factors that act on study subjects or samples, very strict guidelines for the use of transcriptomics in human study subjects have been developed. Typical field conditions associated with doping controls limit the feasibility of following these strict guidelines as there are too many variables counteracting a standardized procedure. After almost a decade of research using transcriptomic tools, it still remains a matter of future technological progress to identify the ultimate biomarker using technologies and/or methodologies that are sufficiently robust against typical biological and technical bias and that are valid in a court of law. Copyright © 2011 John Wiley & Sons, Ltd.

Keywords: doping analysis; transcriptomics; RNA; peripheral blood

Introduction

The fight against doping in sports is significantly dependent on the development of methods and tests for the detection of prohibited substances.[1] Due to the short window of detection for several doping substances, Butch *et al.* suggest that cheating athletes are able to escape detection by training in remote areas, where sample collection is considerably less likely.[2] One of the key points in counteracting the growing problem of doping is the development of more sophisticated tests. More efficient detection with an extended detection window for prohibited substances may prevent athletes from engaging in doping.

Based on the idea that chemical and/or biological doping substances are likely to produce broad genetic, metabolic, and proteomic changes, it has been suggested that all of these doping procedures can be detected by transcriptional profiling.[3] Transcriptional profiling is summarized under the term 'transcriptomics' and generally refers to the characterization and quantification of RNA in specific cells, tissues, or the whole organism. With regard to doping detection, it represents the approach to defining specific biomarkers or expression patterns that indicate the use of prohibited substances or methods. It was hypothesized that even sophisticated methods like gene doping by gene transfer could be detected by transcriptional profiling.[3] Nevertheless, despite the fact that PCR-based direct detection procedures for some specific gene doping applications[4–6] have been recently described, many drugs and doping methods are still undetectable. Therefore, indirect detection approaches using transcriptomics may enable the simultaneous detection of multiple doping substances and procedures. A particularly attractive feature would be that all substances and methods that influence a common transcriptional pathway might be simultaneously detected by using this technique.

In a standardized doping test using a transcriptomic approach, RNA must be obtained from biological samples which should be as homogenous and informative as possible, and furthermore, easily accessible with minimal invasive techniques. As reviewed by Rupert, a number of cell types and body fluids could be used for transcriptional analysis, including blood cells, the hair follicle, buccal and sperm cells, as well as plasma and saliva.[7] Practically, most studies focus on the use of peripheral blood cells, as they are easily accessible and can be obtained in sufficient amounts for gene expression profiling.[8] Investigators from different fields in biomedicine have been screening the transcriptome of peripheral blood cells with the intention of detecting gene expression fingerprints that can be specifically attributed to distinct influences. In medical research, a large number of studies have evaluated gene expression profiles in order to improve diagnostics[9] for diseases like cancer,[10] neurologic disorders,[11] cardiovascular diseases,[12] and asthma,[13] to mention a few only. Furthermore, a number of studies investigated inter- and intra-individual differences in blood cell gene expression related to ethnicity, sex, and age.[14–16]

* *Correspondence to: Professor P. Simon, Department of Sports Medicine, Rehabilitation and Disease Prevention, Johannes Gutenberg-University Mainz, Saarstr. 21, D-55099 Mainz, Germany. E-mail: simonpe@uni-mainz.de*

Department of Sports Medicine, Rehabilitation and Disease Prevention, Johannes Gutenberg-University, Mainz, Germany

As reviewed by Walsh *et al*. the influence of exercise on gene expression using peripheral blood cells has been studied.[17] Aside from this, peripheral blood cell gene expression has been studied in relation to a plethora of other relevant external factors including nutrition,[18] tobacco smoking,[19] and medication use.[20] All of these studies have used chip-based RNA detection technology, which represents one of the core technologies for transcriptomic approaches. It combines standard molecular techniques with high-throughput screening that enables researchers to measure differences in gene expression in a genome-wide manner.

In this review, we will critically discuss the applicability of transcriptomic approaches in doping analysis and research. The first part of the discussion touches on biological variations and highlights individual and external considerations that complicate the interpretation of data from the transcriptome of peripheral blood cells (Figure 1). In the second part, the discussion turns to a number of technical confounding factors that potentially influence the results in the course of sample treatment, sample analysis and data evaluation. The third part presents studies that have utilized transcriptomic approaches for doping detection. Finally, the potentials and pitfalls of transcriptomic approaches in doping analysis and research are discussed.

Biological Variation of the Blood Transcriptome

Peripheral blood presents a convenient source of genetic material. It is easily available in sufficient quantities with minimal invasive techniques, and it represents the most widely used sample type in studies with human subjects. RNA can be easily extracted from distinct blood cells and plasma. However, despite the benefits, blood is a highly complex system with multiple cell types at different stages of their life cycle.[21]

Blood composition

The largest fraction of blood cells is represented by anucleate erythrocytes. Transcriptionally active cells are reticulocytes, leukocytes, and platelets. Leukocytes are nucleated cells that are part of the immune defense system and thereby circulate in blood and lymphatic fluid.[21] Liew *et al*. found that peripheral leukocytes share over 80% of their transcriptome with other human tissues such as brain, colon, heart, kidney, liver, lung, prostate, spleen, and stomach.[22] As a consequence, many studies utilize leukocytes,[23,24] or more specifically, neutrophils[25,26] or peripheral blood mononuclear cells (PBMCs). PBMCs include lymphocytes and monocytes, but not granulocytes or erythrocytes.[27]

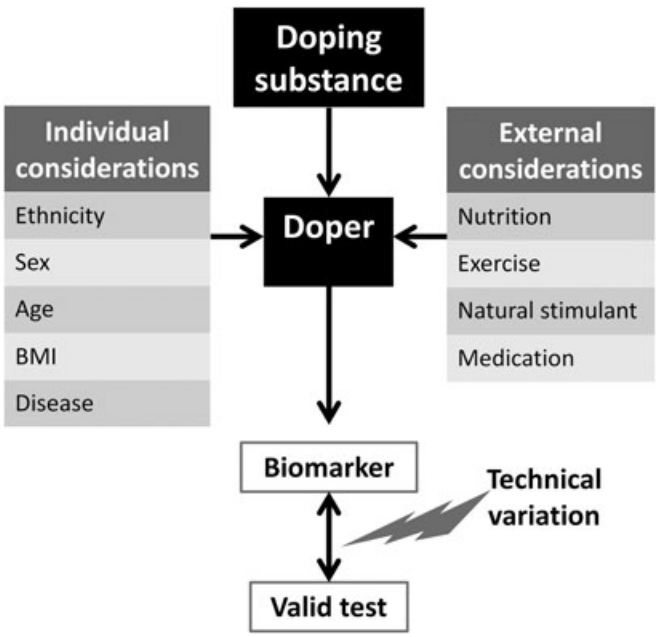

Figure 1. Factors influencing RNA profiling

Individual considerations

In comparison to the large number of studies devoted to examination of differences in gene expression according to external influences and diseases, only a small number of investigators have studied natural intra- and inter-individual[16,28] variation in healthy individuals,[15] as well as the related cell type specific variation in gene expression.[29]

Using a customized 18 000 gene microarray called the 'lymphochip', Whitney *et al*. surveyed the PBMC mRNA expression pattern from 75 healthy individuals (40 male, 35 female).[15] Their study demonstrates that variation in gene expression correlates with age, gender, and the time of day at which samples are collected. The study also suggests that a significant portion of variation in gene expression is related to the shift in the relative proportions of the different peripheral blood cell types.[16] Eady *et al*. found inter-individual variation in PBMC gene expression profiles which correlated with sex, age, and body mass index. An intra-individual comparison showed that the overall expression of many genes is stable within healthy individuals. Approximately 80% of the genes examined had a within subject-variation of less than 20%.[16] However, in another longitudinal study, Radich *et al*. reported large day to day variation in gene expression in peripheral leukocytes.[28] As reviewed by Walsh *et al*. the relative proportion of blood cell types can vary widely even in healthy adults, and disease and exercise further contribute to this effect.[17]

External considerations

A number of studies from various biomedical disciplines have shown alterations of the peripheral blood transcriptome in response to external factors such as nutrition,[30,31] medication,[20] and exercise.[17] For instance, Bouwens *et al*. investigated food supplementation with fish-oil for 26 weeks and found significantly altered gene expression patterns in PBMCs. Using a whole-genome NuGO GeneChip array, 1040 genes were found to be altered in their expression.[18] Lee *et al*. investigated the effects of astemizole, a second-generation antihistamine on heart cells and PBMC gene expression in rats.[20] Using a Rat 230 2.0 array (Affymetrix) Microarray and PBMCs collected 4 h post-treatment, 702 genes were found to be significantly altered in their expression.[20]

Differential gene expression in immune competent cells following exercise has been studied for almost a decade using transcriptomic approaches, and has resulted in a significant amount of knowledge about new pathways and marker molecules in adaptation to exercise. Many studies have used PBMCs to analyze the effects of physical exercise, following the lead of Sonna *et al*.[32] The aim of these studies has been the identification of an exercise-specific 'gene expression fingerprint' that could be used for talent identification, prevention of disease by exercise, and as an additional tool to improve individual/personalized training.[17]

Differential gene expression patterns in response to exercise were found to be dependent on different workloads,[24] gender/age,[33,34] and immunosuppression.[35] Connolly *et al.* were the first investigators to examine the effects of defined exercise on gene expression. Using an Affymetrix HG-U133 GeneChip, a brisk up- and down-regulation of genes associated with stress, inflammation and tissue repair was found. Despite their distinct findings, they could not exclude that changes in PBMC gene-expression profiles after exercise resulted from an altered monocyte to lymphocyte ratio in the circulation.[36] Cellular shifts are a known bias on gene expression patterns and this particular cellular shift depends on the training status of the individual, the intensity of exercise, the nutritional status, as well as the time of the sample collection,[37,38] reviewed Walsh *et al.*[17] In 2007, Büttner *et al.* found that the intensity of the workload influences the level of gene expression.[24] Using the Affymetrix U133 GeneChip, they found that gene expression in leukocytes varied depending on either exhaustive or moderate bouts of exercise. They also reported that changes in gene expression patterns are different when comparing trained and untrained individuals.[24] Northoff *et al.* described that women showed strikingly different patterns of gene regulation in response to physical exercise depending on the phase of their menstrual cycle.[33] Furthermore, Radom-Aizik *et al.* showed that the individual's age influences the gene expression pattern in response to exercise. After ten, 2-min bouts of cycle ergometry, 1320 genes showed an altered expression pattern in early-pubertal girls, whereas 877 genes were changed in late pubertal girls. Importantly, only 622 genes showed a common significant alteration between both groups.[34]

In summary, it is rather difficult to compare the results of these studies as the investigators used different types of exercise, different microarray platforms, different RNA and array preparation methods, different time points of sample collection, and no uniform cell populations. However, in a cross-platform comparison only a small fraction of activated or suppressed genes showed overlap between the different studies that have investigated the supposed effects of exercise on the peripheral blood transcriptome.[39] Although the technology of gene expression profiling in principal enables us to detect the activation of transcriptional pathways, its practical application in exercise physiology and exercise immunology is of very limited and only indicative value at this time.[17]

Technical Variation by Screening the Blood Transcriptome

The partly disappointing results revealed by transcriptomic approaches have focused researchers' attention on the technical confounding factors that may influence the transcriptome in an experimental setting. In transcriptionally active tissues, such as peripheral blood, environmental changes lead to alterations in gene expression patterns.[8] Accordingly, the process of transcriptional profiling itself and the various technical and physical aspects can affect the results of a transcriptomic approach. The entire process includes specimen collection, isolation of cellular components, RNA extraction techniques, and the screening procedure itself.[40]

Sample treatment

The most striking impact on gene expression changes arises from the time delay between sample collection and RNA stabilization and processing.[8,27,40] Using the Affymetrix U133 GeneChip, Debey *et al.* found that a time delay of 20–24 h at room temperature before blood sample preparation led to dramatic changes in PBMC-specific gene expression. These changes pertain to various biological pathways and include genes involved in the immune response, proliferation, metabolism and the cell cycle.[8] Using reverse-transcription quantitative PCR (RT-qPCR), Baecheler et al. found altered gene expression due to short, *ex vivo* incubation times of 1–3 h.[41] Therefore, the handling time for the isolation of specific cell types or sub-cell fractions like PBMCs should be minimized, however this is almost impossible in practice.[8] The use of whole blood RNA can be simplified by the use of RNA stabilization systems. Using PAXGene™ blood can be stored for one day at room temperature, four days at 4 °C and about three months at −20 °C.[40] This allows for transportation of samples over an extended period of time. However, the blood cells are lysed and cannot be sorted or counted. Shifts in the blood cell population that are greater than 2-fold following exercise[36] or more than 5-fold due to different common diseases[21] such as the flu, will severely influence the outcome of such a gene expression analysis even in an intra-individual analysis. Furthermore, even when using PAXGene™ tubes, storage and cycles of freezing/thawing can lead to decreased RNA quality. For instance, Kim *et al.* found a decreased 28S:18S rRNA ratio and lower RNA integrity numbers for RNAs extracted 2–4 d after blood collection in some samples.[40]

Sample processing

Despite the fact that microarrays have undergone technical progress and the utility, scope, and precision of their ability to measure gene expression levels has increased, there are concerns about the reliability and reproducibility of microarray results.[42–45]

In the context of the MicroArray Quality Control project (MAQC), Shi *et al.* tested seven commercially available microarray platforms.[45] One hundred and thirty-seven investigators from 51 organizations measured gene expression levels from two RNA samples in four titration pools on seven microarray platforms. The intention was to test the reproducibility of gene expression revealed by microarrays within a specific test site, across multiple test sites, and the comparability across multiple platforms. Using distinct, high quality RNA samples, this study showed an 89% overlap of differentially expressed genes between test sites when using the same microarray platform, and only 74% across different platforms.[45] In a study by Yang *et al.*, a set of sixteen RNA samples was assayed five times in four different laboratories using a common microarray platform.[43] This study showed great differences in results across laboratories. Their retrospective analysis emphasizes that the identified differences in gene expression resulted from differential specimen treatment. The procedure of the Affymetrix GeneChip array includes labelling, hybridization, wash and staining, and scanning. This study found that different laboratories processed the samples in different batches, which has a substantial impact on microarray data. The resulting batch effects are technical sources of variation that influence gene expression profiles.[43]

Within the MicroArray Quality Control (MAQC)-II study (2010) the investigators investigated the reliability of microarray results to predict different endpoints (e.g. sex of patients and disease progression status). The study involved thiry-six independent teams using six different microarray platforms. A major finding was that proficiency of data analysis could be clearly correlated with the experience of the laboratory team. Irrespective of

experience, some endpoints (e.g. sex of patients) were highly predictable by microarray data, whereas other endpoints (e.g. disease progression status) were difficult to predict.[46]

Despite strict protocols of sample processing and the use of the same microarray platforms within the same laboratory, microarray data show differences. For the validation of microarray results and biomarker monitoring, RT-qPCR remains the method of choice because of its superior sensitivity, reproducibility and dynamic range.[47] Recent technological progress allows for the combination of the sensitivity of classical RT-qPCR with increased throughput. Commercially available microfluidic dynamic arrays allow for measurement of the expression of a large number of genes simultaneously on a single chip.[48] Several studies confirmed high correlation between the results of microfluidic dynamic arrays and RT-qPCR[48–50] with a reproducibility and sensitivity superior to microarray results. However, alterations in gene expression deriving from biological variation and sample treatment and processing remain the same for the various quantitative PCR applications as well as for the other screening approaches in transcriptomics.

Transcriptomic Approaches for Doping Detection

In 2004, Varlet-Marie *et al*.[51] examined gene expression of three selected genes before, during, and after recombinant erythropoietin (rEPO) administration for six weeks in humans using RT-qPCR. They found that the mean mRNA expression slightly increased during rEPO administration, but immediately returned to the initial value following the last injection. Of the three investigated candidates, hemoglobin β, ferritin-light chain, and ornithine decarboxylase antizyme (*OAZ*), only *OAZ*-expression increased in all subjects, presenting itself as a potential subject-independent marker. Unfortunately, this gene was also affected by the withdrawal of iron supplementation in the control group. In 2009, Varlet-Marie *et al*.[52] constructed three Serial Analysis of Gene Expression (SAGE) libraries from pooled blood samples of healthy recreational athletes after subcutaneous treatment with rEPO ($0.72\,\mu g\,kg^{-1}wk^{-1}$) for four weeks. Ninety five genes were found to be differentially expressed. Using RT-qPCR, these genes were tested in two athletes after high dose rEPO administration (~260 IU/kg) for 11 days followed by low-dose (<10% of initial dose) rEPO treatment for another 22 days. Thirty three genes were differentially expressed during high dose rEPO administration. When data for high- and low-dose administration were combined, five genes were significantly altered for at least one week post administration in both subjects. However, gene expression varied widely over time with all genes showing a transient return to baseline levels following withdrawal of the high dose and during the mirco-dose administration phase.

In 2010, Bailly-Chouriberry *et al*.[53] examined gene expression profiles of seven thoroughbred horses after subcutaneous injection of rEPO (40 IU/kg/day) on six consecutive days. For biomarker identification, they constructed three SAGE libraries using pooled blood samples from seven horses collected before, during, and after rEPO administration. Analysis of pooled blood samples from the horses in subsequent RT-qPCR analysis over time indicated that there might be eight candidates that are likely to respond transcriptionally during rEPO administration. However, the candidate genes revealed in this study were not verified on an individual basis, and none of the candidates were found in the human study by Varlet-Marie *et al*.[52]

Mitchell *et al*.[54] conducted a very well-designed study with the objective to detect GH doping in humans. Considering that growth hormone (GH) regulates various components of the immune system and affects the activity and gene expression of various immune-competent cells, they hypothesized that GH treatment may have effects on the peripheral blood transcriptome. Therefore, 7 male and 13 female recreational athletes received supra-physiological doses (2 mg/d) of GH for 8 weeks. Using the Agilent microarray platform, this study showed significant gene expression changes compared to the control group for 41 candidate-genes in men and 353 candidate-genes in women: Surprisingly, none of those differentially expressed genes were found to be common between men and women. In contrast to this, Insulin Like Growth Factor 1 protein levels could be found substantially elevated in all GH treated study subjects throughout the treatment period, indicating the effectiveness of GH administration. Taken together, the effect of GH treatment led to very low expression changes within the range of normal variation in inter-individual gene expression. Hence, discriminant gene signatures could not be defined and the authors conclude that transcriptional profiling of mRNA from PBMCs does not seem to be a viable approach for the detection of GH doping. Mitchell *et al*.[54] note that some effects may be masked by the mixed cell population and the data might reflect a net of multiple, possibly opposing effects on the same transcript in different sub-populations.

In 2010, Minella *et al*.[55] developed a low density microarray platform, called 'AndroChip 2' for putative detection of anabolic steroid abuse. The microarray contains 190 probes to detect gene expression profiles of androgen and insulin specific pathways. In order to determine whether gene expression is dependent on sports activity, they determined the gene expression profile of mRNA pools from athletes having daily intense training (n=20), athletes who practice sports occasionally (n=19), and sedentary subjects (n=15). Microarray results and subsequent RT-qPCR validation for 19 genes in 30 subjects emphasize that the mean mRNA expression level does not differ between the three groups. Whether this low density microarray platform can provide valid indications of androgen abuse remains to be demonstrated.

Potential and Pitfalls of Transcriptomic Approaches

Because some doping substances share a similar or identical structure with naturally occurring endogenous substances, their detection has been markedly reduced or impossible thus far. During the past few years, special attention was directed towards improving the detectability of anabolic agents, GH, erythropoiesis-stimulating agents (ESAs), and gene doping.[56] Since ESAs were developed for the treatment of anemia, they have been widely abused in sports to improve oxygen transport capacity, and therefore to increase endurance performance. An argument for transcriptomic approaches or gene expression analysis can be particularly well elucidated for ESAs. ESAs are a very heterogenous group of substances which include rEPO variants, EPO biosimilars and analogues, EPO mimetics, as well as gene doping through delivery of the EPO gene in a constitutively active or inducible form.[57] Despite this considerable variation, all ESAs converge on the same downstream-mechanism through activation of the EPO receptor.[58] A particular advantage of transcriptional profiling could be the identification of potential biomarkers or gene expression profiles that specifically detect

activation of the common pathway downstream of EPO-receptor activation .[52] Such an indirect detection technique will have to deliver results that are robust, valid, sensitive, and highly specific, such that they could be used as a robust tool in the legal justification of sanctions against athletes.

Indirect markers for doping detection have been used for many years. According to the review by Sottas,[59] early approaches relied on population-based reference intervals, thereby ignoring individual differences. Currently, indirect detection approaches allow for the combination of population-derived limits with individual baselines for certain markers. Using Bayesian statistics, a set of internal and external factors such as age, gender, and exposure to altitude are considered in determining a subject's specific threshold indicative for doping. Bayesian networks related to blood doping or anabolic steroid abuse have been applied and validated using more than 20 000 blood or steroid profiles.[60] Following preliminary discussions in 2002, the World Anti-Doping Agency (WADA) began development of the Athletes Biological Passport (ABP) in 2006.[61] Three modules are currently identified: hematological-, steroidal-, and endocrinological. So far, only the hematological module has been implemented. However, additional variables derived from a 'transcriptome module' could likely be added. Additionally, potential mRNA biomarkers for ESAs, as suggested by Varlet-Marie *et al*.[52] could in principal be integrated into the hematological module. However, a large number of samples must be tested in order to determine the variation in gene expression according to biological variables such as sex, age,[16] ethnicity,[14] as well as external variables such as exercise,[24] nutrition,[18] and medication,[20] and finally according to technical variation related to sample retrieval, storage and processing.[8]

mRNA expression is highly associated with epigenetic factors such as DNA methylation and chromatin modification. These epigenetic signatures are greatly cell-type specific but furthermore are also influenced by environmental factors such as life experiences,[62] physical activity,[63] age and environment,[64] and nutrition[65] to mention a few only. Additionally, gene expression patterns can be remarkably influenced by the subject's underlying genetic sequence which differs considerably within the individual and across ethnic groups. To date, 11 million single nucleotide polymorphisms (SNPs) have been identified in the human genome thus far, and many more are expected to be identified in the future.[66] Potential effects on gene expression of copy number variations (CNVs), rare variations, and mutations should also be taken into account.[67] Such biological variations also act as confounding factors in studies that aim to reveal specific gene expression signatures; therefore the nature of their effect on gene expression has to be determined before this technique can contribute to the field of doping analysis. In this context it is noteworthy that SNPs can lead to biological variation in mRNA transcript abundance, and can also lead to technical variation in measured transcript abundance. Gene chip probes are designed and optimized to target reference sequences of a particular genome and the binding efficiency of probes has been shown to be influenced by sequence variation.[68]

Screening for biomarkers using microarrays or other transcriptomic techniques is therefore highly prone to delivering false positive results that have to be validated in large-scale confirmation studies. It is therefore important to statistically demonstrate whether a potential biomarker is sufficiently robust, at least on an intra-individual level. When compared to the naturally occurring intra-individual variation in gene expression and technical variation, the small gene expression changes in response to doping found by Varlet-Marie *et al*.[52] and Mitchell *et al*.[54] are a limiting factor in the integration of transcriptomic analysis in doping analysis. As has been previously shown, PBMC-derived gene expression is highly influenced by the time of day, exercise, and even food intake. Thereby, many clinical studies have emphasized the need to control for such external confounding factors.[9] This introduces the requirement for subjects to fast before blood sampling. Such standardization procedures are not possible in an elite athlete population. In 2009, WADA published the *Athletes Biological Passport Operation Guidelines*.[69] These guidelines were established to standardize the results of monitored variables. Accordingly, blood must be sampled >2 h after training or competition, and should be analyzed within 36 h. To ensure test accuracy, the additional conditions that may effect gene expression must be assessed and considered in the ABP. For professional athletes who are tested under the typical field conditions of a doping test, this will be by far more laborious than the efforts that have been made to establish the blood passport module. For instance, exercise must be controlled for, since exercise can be followed by a significant gene expression response for up to 24 h.[70] A major pitfall could be gene expression changes in response to cellular shifts. As reviewed in Walsh *et al*., an acute bout of exercise is followed by an increase in blood neutrophil levels for several hours, a transient monocytosis (~2 h), and a 50–100% increase in natural killer cells in PBMCs.[17] Changes in whole blood or PBMC-specific gene expression cannot be distinguished from cellular shift dependent gene expression.[17] Additionally, some effects of the doping substance may be masked due to opposing effects on different blood cell types.[54] Thus, another critical factor for the intra-individual comparison of gene expression is the given blood cell population that can vary not only due to training and competition, but also due to minor infections,[71] and the menstrual cycle.[33] Other factors, such as the ambient temperature at which a sample was taken, and the time interval between sample storage and processing, have also been shown to be critical and must be carefully controlled. Given the high influence of storage conditions on gene expression analysis, only the use of PAXGene™ blood seems to be acceptable. This means that the blood samples that are currently used in the blood pass system cannot be used to establish or evaluate biomarker candidates. For this purpose, future blood samples drawn from the elite athlete population have to be taken following strict guidelines warranting mRNA stability.

The current blood pass system is continuously monitored by experts on an executive level in order to determine whether a potential sanction seems justified given the profile changes observed. It is interesting to consider whether such experts could be found for the interpretation of transcriptomic profiles. The current blood passport is based on data that has been collected for over half a century, on a frequent basis, in hundreds of millions of people worldwide. In addition, several thousand studies have already demonstrated how particular cell types of the hematopoietic system react to various influences.

At this time, knowledge of the transient variables influencing RNA data is still too limited for the implementation of transcriptomics into a biological passport. Additionally, we are neither able to define 'normal' gene expression under various conditions nor to differentiate data which are influenced by technical variations.

Thus far, the heterogeneity and incongruence of the outcomes of studies that either attempted to find potential biomarkers for

exercise or potential biomarkers for certain doping substances does not justify the conclusion that the principal problems elucidated here can be overcome easily in the near future.

Novel methodologies, such as next geneneration sequencing may lead to the identification of potential biomarkers that can be used in the future. However, if very robust, consistent, and significant expression or profile differences cannot be revealed in a typical pilot setting using such novel, groundbreaking technology, we will have to keep in mind that our knowledge of the transcriptome, its regulation, and its confounding variables, is still very limited.

Until transcriptomics can be implemented as a standard analytical tool in doping detection, we must collect more data and gain additional experience to ensure that the interpretation of results leads to yes/no messages, without any doubt.

As the blood transcriptome is a particularly dynamic and susceptible system, the use of mRNA from other tissues such as the hair follicle[72,73] also should be considered. Reiter *et al.* studied gene expression changes of cultured hair follicle dermal papilla cells in response to stanozolol treatment and showed significant gene expression changes for selected target genes. The use of hair papilla cells would reduce some types of bias, such as cellular shifts.[73]

Outlook: Other RNAs as potential biomarkers

When the human genomic sequence was published in 2004, it came as a surprise that only approximately 1.2% of the entire genomic DNA codes for protein.[74] In this context, it is noteworthy that during the past few decades, non-coding RNAs (ncRNAs) were identified and have been investigated in further detail. Functionally, most of them are involved in the processes of the translational machinery. In 1993, a novel class of small regulatory RNAs, the so-called microRNAs (miRNA) were identified.[75] miRNAs are short (~22 nucleotides) highly conserved non-coding RNA molecules that post-transcriptionally modulate the translation of numerous proteins by binding to the 3′ or 5′ untranslated regions (UTRs) of their target mRNAs.[76] Physiologically, miRNAs are involved in the regulation of cellular processes such as cell differentiation, metabolism, and neural development in a highly cell-type specific manner.[77] Interestingly miR-1, miR-133a, and miR-206 where found to be solely expressed in muscle.[78] A genetic link between muscle hypertrophy and miRNAs has been observed in Texel sheep. A G/A mutation in the 3′UTR of the myostatin gene which is involved in the repression of muscle growth was found to increase the binding efficiency for miR-1 and miR-206, leading to suppressed protein translation.[79] It has been proposed that differential miRNA expression might serve as a novel biomarker for a magnitude of traits[80] and it cannot be excluded that miRNAs could also serve as a helpful tool in the detection of doping. Interestingly, the expression of miR-188, miR-362 and miR-210 has already been shown to be significantly up-regulated after EPO stimulation in UT-7 cells.[81]

In addition to their intracellular role, miRNAs are also detectable in serum. Secreted in exosomes, they are highly stable and have already been used as circulating biomarkers in cancer,[82] diabetes,[83] and myo-cardial infarction.[84] In this context, it is noteworthy that the first highly promising research article was recently published, focusing on small interfering RNA using plasma samples for doping control purposes on a mass spectrometric platform.[85] Interestingly, since the funding of a research project by WADA in 2010, miRNAs are investigated as potential biomarkers for the detection of erythropoiesis stimulating agent abuse.[86] Further research is needed to clarify the reliability of the information coming from circulating ncRNAs as indicators of certain traits and diseases and whether they can also be used as a doping-related biomarker.

However, analogous to the problems already discussed for mRNA profiling, similar biological and technical problems can be expected for non-coding RNAs as cellular biomarkers used in a doping test scenario.

Conclusion

Human blood transcriptome screening has been used widely in various fields, including doping detection trials. The goal with respect to doping detection is to discover specific gene expression changes that allow indirect detection of the abuse of prohibited substances. The studies emphasize that transcriptional profiling does not provide a clear 'yes' or 'no' answer in response to doping abuse. A particular problem of most biomarkers, including RNAs, is the significant inter-individual variation between population, sex, and age, which is furthermore influenced by a large number of other confounding factors such as exercise, nutrition, and medication. Transcriptional profiling can be implemented into the ABP as an indirect detection approach upon adequate control of internal and external confounding factors. In order to accomplish this goal, a large number of samples would have to be tested to determine variation in gene expression according to biological-, external-, and technical factors. As discussed before, the blood transcriptome is a particularly dynamic and susceptible system that may hamper the predictive power for doping detection. Therefore, other RNA sample sources should be considered. A major problem is that biological and technical variations do not only coexist, but also interact in many ways. Theoretically, randomization procedures as well as strict experimental settings should lead to a minimization of most confounding factors that influence the peripheral blood transcriptome. However, such standardized conditions cannot be achieved in typical doping control situations. As technologies continue to improve, the identification of specific transcriptional pathways and potential biomarkers cannot be excluded. At present, the use of transcriptomic approaches or potential biomarkers revealed by transcriptomics is unlikely to be sufficiently robust to withstand biological and technical bias.

Acknowledgement

The authors thank Magdalena Jurkiewicz for her critical reading of the manuscript.

References

[1] V. Ivanova, T. Boghosian, O. Rabin. *Accredit. Qual. Assur.* **2007**; *12*, 491.
[2] A. W. Butch, J. A. Lombardo, L. D. Bowers, J. Chu, D. A. Cowan. *Clin. Chem.* **2011**; *57*, 000.
[3] T. Friedmann, O. Rabin, M. S. Frankel. *Science* **2010**; *327*, 647.
[4] T. Beiter, M. Zimmermann, A. Fragasso, S. Armeanu, U. M. Lauer, M. Bitzer, H. Su, W. L. Young, A. M. Niess, P. Simon. *Exerc. Immunol. Rev.* **2008**; *14*, 73.
[5] T. Beiter, M. Zimmermann, A. Fragasso, J. Hudemann, A. M. Niess, M. Bitzer, U. M. Lauer, P. Simon. *Gene Ther.* **2010**; *18*, 225.
[6] W. Ni, C. Le Guiner, G. Gernoux, M. Penaud-Budloo, P. Moullier, R. O. Snyder. *Gene Ther.* **2011**; *1*.

[7] J. L. Rupert. *Scand. J. Med. Sci. Sports* **2009**; *19*, 753.
[8] S. Debey, U. Schoenbeck, M. Hellmich, B. S. Gathof, R. Pillai, T. Zander, J. L. Schultze. *Pharmacogenomics J.* **2004**; *4*, 193.
[9] S. Mohr, C. C. Liew. *Trends Mol. Med.* **2007**; *13*, 422.
[10] P. Sharma, N. S. Sahni, R. Tibshirani, P. Skaane, P. Urdal, H. Berghagen, M. Jensen, L. Kristiansen, C. Moen, A. Zaka. *Breast Cancer Res.* **2005**; *7*, R634.
[11] O. C. Maes, S. Xu, B. Yu, H. M. Chertkow, E. Wang, H. M. Schipper. *Neurobiol. Aging* **2007**; *28*, 1795.
[12] J. Ma, C.C. Liew. *J. Mol. Cell. Cardiol.* **2003**; *35*, 993.
[13] N. N. Hansel, S. C. Hilmer, S. N. Georas, L. M. Cope, J. Guo, R. A. Irizarry, G. B. Diette. *J. Lab. Clin. Med.* **2005**; *145*, 263.
[14] J. D. Storey. *Am. J. Hum. Genet.* **2007**; *80*, 1194.
[15] A. R. Whitney, M. Diehn, S. J. Popper, A. A. Alizadeh, J. C. Boldrick, D. A. Relman, P. O. Brown. *Proc. Natl. Acad. Sci. USA* **2003**; *100*, 1896.
[16] J. J. Eady, G. M. Wortley, Y. M. Wormstone, J. C. Hughes, S. B. Astley, R. J. Foxall, J. F. Doleman, R. M. Elliott. *Physiol. Genomics* **2005**; *22*, 402.
[17] N. P. Walsh, M. Gleeson, R. J. Shephard, J. A. Woods, N. C. Bishop, M. Fleshner, C. Green, B. K. Pedersen, L. Hoffman-Goetz, C. J. Rogers. *Exerc. Immunol. Rev.* **2011**; *17*, 6.
[18] M. Bouwens, O. van de Rest, N. Dellschaft, M. G. Bromhaar, L. C. de Groot, J. M. Geleijnse, M. Muller, L. A. Afman. *Am. J. Clin. Nutr.* **2009**; *90*, 415.
[19] D. M. van Leeuwen, R. W. Gottschalk, M. H. van Herwijnen, E. J. Moonen, J. C. Kleinjans, J. H. van Delft. *Toxicol. Sci.* **2005**; *86*, 200.
[20] E. H. Lee, J. H. Oh, H. J. Park, D. G. Kim, J. H. Lee, C. Y. Kim, M. S. Kwon, S. Yoon. *Arch. Toxicol.* **2010**; *84*, 609.
[21] A. V. Hoffbrand, P. A. H. Moss, J. E. Pettit. Essential haematology, Vol. *5th ed*, Blackwell Science Publications, Oxford, **2006**.
[22] C. C. Liew, J. Ma, H. C. Tang, R. Zheng, A. A. Dempsey. *J. Lab. Clin. Med.* **2006**; *147*, 126.
[23] T. Hilberg, H. P. Deigner, E. Moller, R. A. Claus, A. Ruryk, D. Glaser, J. Landre, F. M. Brunkhorst, K. Reinhart, H. H. Gabriel. *FASEB J.* **2005**; *19*, 1492.
[24] P. Büttner, S. Mosig, A. Lechtermann, H. Funke, F. C. Mooren. *J. Appl. Physiol.* **2007**; *102*, 26.
[25] S. Radom-Aizik, F. Zaldivar, Jr., S. Y. Leu, P. Galassetti, D. M. Cooper. *J. Appl. Physiol.* **2008**; *104*, 236.
[26] S. Radom-Aizik, F. Zaldivar Jr., S. Oliver, P. Galassetti, D. M. Cooper. *J. Appl. Physiol.* **2010**; *109*, 252.
[27] H. Fan, P. S. Hegde. *Curr. Mol. Med.* **2005**; *5*, 3.
[28] J. P. Radich, M. Mao, S. Stepaniants, M. Biery, J. Castle, T. Ward, G. Schimmack, S. Kobayashi, M. Carleton, J. Lampe. *Genomics* **2004**; *83*, 980.
[29] C. Palmer, M. Diehn, A. A. Alizadeh, P. O. Brown. *BMC Genomics* **2006**; *7*, 115.
[30] L. Afman, M. Muller. *J. Am. Diet. Assoc.* **2006**; *106*, 569.
[31] M. Kussmann, L. Krause, W. Siffert. *Nutr. Rev.* **2010**; *68* Suppl 1, S38.
[32] L. A. Sonna, C. B. Wenger, S. Flinn, H. K. Sheldon, M. N. Sawka, C. M. Lilly. *J. Appl. Physiol.* **2004**; *96*, 1943.
[33] H. Northoff, S. Symons, D. Zieker, E. V. Schaible, K. Schafer, S. Thoma, M. Loffler, A. Abbasi, P. Simon, A. M. Niess. *Exerc. Immunol. Rev.* **2008**; *14*, 86.
[34] S. Radom-Aizik, F. Zaldivar, Jr., S. Y. Leu, D. M. Cooper. *J. Appl. Physiol.* **2009**; *107*, 168.
[35] I. Konigsrainer, D. Zieker, M. Loffler, S. Buhler, M. Walter, S. Beckert, J. Glatzle, H. Northoff, S. Nadalin, A. Konigsrainer. *Exerc. Immunol. Rev.* **2010**; *16*, 184.
[36] P. H. Connolly, V. J. Caiozzo, F. Zaldivar, D. Nemet, J. Larson, S. P. Hung, J. D. Heck, G. W. Hatfield, D. M. Cooper. *J. Appl. Physiol.* **2004**; *97*, 1461.
[37] D. Zieker, E. Fehrenbach, J. Dietzsch, J. Fliegner, M. Waidmann, K. Nieselt, P. Gebicke-Haerter, R. Spanagel, P. Simon, A. M. Niess. *Physiol. Genomics* **2005**; *23*, 287.
[38] S. Nakamura, M. Kobayashi, T. Sugino, O. Kajimoto, R. Matoba, K. Matsubara. *Biochem. Bioph. Res. Co.* **2010**; *391*, 846.
[39] E. Fehrenbach. *J. Appl. Physiol.* **2007**; *102*, 7.
[40] S. J. Kim, D. J. Dix, K. E. Thompson, R. N. Murrell, J. E. Schmid, J. E. Gallagher, J. C. Rockett. *Clin. Chem.* **2007**; *53*, 1038.
[41] E. C. Baechler, F. M. Batliwalla, G. Karypis, P. M. Gaffney, K. Moser, W. A. Ortmann, K. J. Espe, S. Balasubramanian, K. M. Hughes, J. P. Chan. *Genes Immun.* **2004**; *5*, 347.
[42] P. Jaluria, K. Konstantopoulos, M. Betenbaugh. J. Shiloach, *Microb. Cell Fact.* **2007**; *6*, 4.
[43] H. Yang, C.A. Harrington, K. Vartanian, C. D. Coldren, R. Hall, G. A. Churchill. *PLoS One* **2008**; *3*, e3724.
[44] C. King, N. Guo, G. M. Frampton, N. P. Gerry, M. E. Lenburg, C. L. Rosenberg. *J. Mol. Diagn.* **2005**; *7*, 57.
[45] L. Shi, L.H. Reid, W. D. Jones, R. Shippy, J. A. Warrington, S. C. Baker, P. J. Collins, F. de Longueville, E. S. Kawasaki, K. Y. Lee. *Nat. Biotechnol.* **2006**; *24*, 1151.
[46] L. M. Shi, G. Campbell, W. D. Jones, F. Campagne, Z. N. Wen, S. J. Walker, Z. Q. Su, T. M. Chu, F. M. Goodsaid, L. Pusztai. *Nat. Biotechnol.* **2010**; *28*, 827.
[47] H. D. VanGuilder, K. E. Vrana, W. M. Freeman. *Biotechniques* **2008**; *44*, 619.
[48] S. L. Spurgeon, R. C. Jones, R. Ramakrishnan. *PLoS One* **2008**; *3*, e1662.
[49] B. Lu, J. Xu, J. Chen, J. Yu, E. Xu, M. Lai. *Clin. Chim. Acta* **2008**; *389*, 146.
[50] J. S. Jang, V. A. Simon, R. M. Feddersen, F. Rakhshan, D. A. Schultz, M. A. Zschunke, W. L. Lingle, C. P. Kolbert, J. Jen. *BMC Genomics* **2011**; *12*, 144.
[51] E. Varlet-Marie, M. Audran, M. Lejeune, B. Bonafoux, M. T. Sicart, J. Marti, D. Piquemal, T. Commes. *Haematologica* **2004**; *89*, 991.
[52] E. Varlet-Marie, M. Audran, M. Ashenden, M. T. Sicart, D. Piquemal. *Am. J. Hematol.* **2009**; *84*, 755.
[53] L. Bailly-Chouriberry, F. Noguier, L. Manchon, D. Piquemal, P. Garcia, M. A. Popot, Y. Bonnaire. *Drug Test. Analysis* **2010**; *2*, 339.
[54] C. J. Mitchell, A. E. Nelson, M. J. Cowley, W. Kaplan, G. Stone, S. K. Sutton, A. Lau, C. M. Lee, K. K. Ho. *J. Clin. Endocrinol. Metab.* **2009**; *94*, 4703.
[55] D. Minella, M. Biancolella, B. Testa, G. Prosperini, R. Zenobi, G. Novelli, M. G. Giganti. *J. Biol. Regul. Homeost. Agents* **2010**; *24*, 413.
[56] M. Thevis, T. Kuuranne, H. Geyer, W. Schanzer. *Drug Test. Anal* **2011**; *3*, 1.
[57] W. Jelkmann. *Curr. Med. Chem.* **2009**; *16*, 1236.
[58] S. Elliott, E. Pham, I.C. Macdougall. *Exp. Hematol.* **2008**; *36*, 1573.
[59] P. E. Sottas, M. Saugy, C. Saudan. *Endocrinol. Metab. Clin. North Am.* **2011**; *39*, 59.
[60] Swiss Laboratory for Doping Analysis: Athletes Biological Passport. Available at: http://www.doping.chuv.ch/en/lad_home/lad-recherche-developpement/lad-recherche-developpement-projets-finalises/lad-recherche-developpement-projets-finalises-pass-biol.htm [14 June **2011**].
[61] WADA: Athletes Biological Passport. Available at: http://www.wada-ama.org/en/Resources/Q-and-A/Athlete-Passport/ [16 June **2011**].
[62] I. C. Weaver, *Epigenetics* **2007**; *2*, 22.
[63] F. F. Zhang, R. Cardarelli, J. Carroll, S. Zhang, K. G. Fulda, K. Gonzalez, J. K. Vishwanatha, A. Morabia, R. M. Santella. *Epigenetics* **2011**; *6*, 293.
[64] M. F. Fraga, E. Ballestar, M. F. Paz, S. Ropero, F. Setien, M. L. Ballestar, D. Heine-Suner, J. C. Cigudosa, M. Urioste, J. Benitez. *Proc. Natl. Acad. Sci. USA* **2005**; *102*, 10604.
[65] R. A. Waterland, R. L. Jirtle. *Mol. Cell. Biol.* **2003**; *23*, 5293.
[66] J. M. Kidd, G. M. Cooper, W. F. Donahue, H. S. Hayden, N. Sampas, T. Graves, N. Hansen, B. Teague, C. Alkan, F. Antonacci. *Nature* **2008**; *453*, 56.
[67] E. Chaignat, E. A. Yahya-Graison, C. N. Henrichsen, J. Chrast, F. Schutz, S. Pradervand, A. Reymond. *Genome Res.* **2011**; *21*, 106.
[68] D. C. Ciobanu, L. Lu, K. Mozhui, X. S. Wang, M. Jagalur, J. A. Morris, W. L. Taylor, K. Dietz, P. Simon, R. W. Williams. *Genetics* **2010**; *184*, 119.
[69] WADA: Athletes Biological Passport Operation Guidelines. Available at: http://www.wada-ama.org/Documents/Resources/Guidelines/WADA_ABP_OperatingGuidelines_EN_2.1.pdf [16 June **2011**].
[70] T. Kawai, K. Morita, K. Masuda, K. Nishida, A. Sekiyama, S. Teshima-Kondo, Y. Nakaya, M. Ohta, T. Saito, K. Rokutan. *Clin. J. Sport Med.* **2007**; *17*, 375.
[71] W. Stock, R. Hoffman. *Lancet* **2000**; *355*, 1351.
[72] M. Reiter, S. Lüderwald, M.W. Pfaffl, H. H. Meyer. *Doping J.* **2008**; *5*.
[73] M. Reiter, M. W. Pfaffl, M. Schonfelder, H. H. Meyer. *Biomark Insights* **2008**; *4*, 1.
[74] Z. Abdellah, A. Ahmadi, S. Ahmed, M. Aimable, R. Ainscough, F. Almeida, C. Almond, A. Ambler, K. Ambrose, K. Ambrose. *Nature* **2004**; *431*, 931.
[75] R. C. Lee, R. L. Feinbaum, V. Ambros. *Cell* **1993**; *75*, 843.
[76] D. P. Bartel *Cell* **2004**; *116*, 281.
[77] W. P. Kloosterman, R. H. Plasterk *Dev. Cell* **2006**; *11*, 441.
[78] E. J. Lee, M. Baek, Y. Gusev, D.J. Brackett, G. J. Nuovo, T. D. Schmittgen. *Rna –A Publ of the Rna Soc* **2008**; *14*, 35.

[79] M. Thomas, B. Langley, C. Berry, M. Sharma, S. Kirk, J. Bass, R. Kambadur *J. Biol. Chem.* **2000**; *275*, 40235.
[80] H. M. Heneghan, N. Miller, M. J. Kerin. *Curr. Opin. Pharmacol.* **2010**; *10*, 543.
[81] N. Kosaka, K. Sugiura, Y. Yamamoto, Y. Yoshioka, H. Miyazaki, N. Komatsu, T. Ochiya, T. Kato. *Br. J. Haematol.* **2008**; *142*, 293.
[82] E. K. Ng, W. W. Chong, H. Jin, E. K. Lam, V. Y. Shin, J. Yu, T. C. Poon, S. S. Ng, J. J. Sung. *Gut* **2009**; *58*, 1375.
[83] C. Guay, E. Roggli, V. Nesca, C. Jacovetti, R. Regazzi *Transl. Res.* **2011**; *157*, 253.
[84] G. W. Dorn 2nd, *Transl. Res.* **2011**; *157*(4): 226–35.
[85] M. Kohler, A. Thomas, K. Walpurgis, W. Schanzer, M. Thevis *Anal. Bioanal. Chem.* **2010**; *398*, 1305.
[86] WADA: Funded Research Projects. Available at: http://www.wada-ama.org/en/Science-Medicine/Research/Funded-Research-Projects/ [17 June **2011**].