18 Biomarker Discovery via RT-qPCR and Bioinformatical Validation

Christiane Becker, Irmgard Riedmaier, and Michael W. Pfaffl

CONTENTS

18.1 INTRODUCTION

There is a growing interest in life science research in the use of expressed transcripts that form the basis of biological markers (biomarkers) and in addressing some of the challenging statistical issues that arise when attempting to validate them. Biomarkers have extensively been used across diagnostic and therapeutic areas of many life science disciplines, including clinical, physiological, biochemical, developmental, morphological, and molecular applications.[1] Biomarkers have been defined as "cellular, biochemical or molecular alterations that are measurable in biological media such as human tissues, cells, or fluids."[2] The official definition, developed by the "Biomarkers definitions working group" of the NIH is[3]: "A biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention." More recently the definition has been broadened to include more biological characteristics that can be objectively measured and evaluated as a biological indicator.[4] A biomarker can refer to any measurable molecular, biochemical, cellular, or morphological alternations in biological media such as human tissues, cells, or fluids.[5]

18.1.1 Biomarkers at Various Molecular Levels

Advances in genomics, proteomics, transcriptomics, and metabolomics have generated many candidate biomarkers with the potential for diagnostic and clinical value. Current efforts are focused on biomarker discovery, reliable detection, and early diagnosis, for example, in cancer biology through the application of various–omics technologies. The success of biomarker identification depends on

many factors, such as the type of molecule (e.g., gene, transcript, protein, metabolite), the intensive validation across a heterogeneous population and its variations (age, sex, species, breed), the quality and integrity of the biological sample, the size of the dataset(s) used, and the statistical methods that were applied for validation. Probably the most significant factor leading to success is the number of variables and conditions being tested, because, what appears to be specific in a given biological dataset may not necessarily be so in a larger set or even in the entire population. Therefore, the more conditions and variables being evaluated; the better will be the outcome of the prediction and the validity of the discovered biomarkers.[6,7]

The integration of different technologies on various–omics levels for data collection and their analyses are pivotal for biomarker identification, characterization, validation, and successful usage. The application of integrative functional informatics represents a novel direction in such biomarker discovery and brings a new dimension to molecular diagnostics.[1] These markers can represent the combination of multiple pieces of information on various biological levels, such as genes, their mutations, SNPs, gene methylation pattern, alternative gene transcripts (mRNA and miRNA), posttranslational modified proteins, metabolites, morphological changes, or altered physiological responses.

The first step in all biological and physiological processes is the transcription of specific genes into mRNAs and noncoding RNAs as prerequisite for the generation of functional proteins. Gene expression is a dynamic process that adapts rapidly to physiological changes or exogenous stimuli and thus the transcriptome with its enormous number of alternative spliced mRNAs, large and small noncoding RNAs reflects the current physiological situation in different tissues, organs, and even in single cells.[8] Therefore, monitoring the transcriptome is, potentially, a very promising approach for detecting biomarkers for specific physiological situations, diseases, or treatments.

Further biomarkers can be discovered at the level of the proteome and the metabolome. To investigate the complex proteome, applied proteomic technologies are used to separate, identify, and characterize a global set of proteins. In addition, information should be provided about the protein concentration, tissue, or cellular location, any modifications or functional attachments, and interactions, for example, protein–protein, protein–DNA, or protein–ligand.[9] The proteome, unlike the "fixed" genome, possesses an intrinsic complexity and is in a constant state of flux. The benefit of protein analysis is the ability to take into account posttranslational modifications, which can markedly alter the function, activity or half-life of a protein. In addition, the final amount of the active protein can differ greatly from the initial amount of mRNA transcribed and present in the cell.

Metabolomics is a relatively new discipline that can facilitate rapid *in vivo* screening of various factors, including drug efficacy and/or toxicity and underlying physiological processes. The metabolomic approach is complementary to the other–omic profiling technologies and can provide a chemical and biochemical profile of a specific body fluid, organ, or tissue during a continuous time-course analysis.[9] Overall, metabolomics can facilitate the determination of metabolic profiles and the mapping of interactions between metabolic pathways across organisms.

As described, there are various ways to discover biomarkers: at the level of the genome, transcriptome, proteome, or metabolome. Herein, we will focus on biomarkers identified in the transcriptome at the RNA level, with the current focus on mRNA and miRNA. Each gene has its set of characteristic expression profiles and alternative splice variants, that is, in which cells or tissues and at what time it is expressed and how it responds to environmental stimuli. For research and biomedical purposes, only a few genes may be sufficiently reliable to be used as indicators of healthy or diseased biological states.[7] Approaches centered on transcriptomics consist of various methods to measure the expression of genes, including microarray analysis, RT-PCR-based methods or holistic assumption-free methods such as next-generation sequencing technologies.[10,11] Quantitative RT-PCR and microarray-based analysis have significantly expanded the throughput of expression studies, and numerous examples of potential microarray-based biomarkers have been published.[12–14]

18.1.2 Biomarkers at the mRNA Level

There are different methods available to quantify single transcripts. In general, these methods differ in the number of quantifiable genes. Northern blot is the classical approach for the detection of different mRNAs. A more precise method for the quantification of gene expression is quantitative real-time RT-PCR (RT-qPCR). With both methods it is possible to quantify single genes or multiple gene sets in one run. Using qPCR arrays, up to 384 different transcripts can be analyzed in parallel. There are also other methods available for a holistic screen of gene-expression changes. Until recently, microarray analysis has been the screening method of choice for most gene-expression experiments at the mRNA and miRNA level. With this approach, a sample can be screened for the expression of all known transcripts present in the gene database. But how can new, unknown genes, alternative splice variants, or miRNA intermediates be measured? RNA sequencing (RNA-Seq) is a new method which permits the sequencing as well as the quantification of the whole transcriptome of a biological sample. It is a very sensitive approach; a single transcript of a given gene is detectable, and since it is assumption-free it is also possible to discover new transcripts or unknown splice variants.[10,11]

The application of transcriptomics to biomarker research has successfully been used in various fields of life science. In molecular medicine, it has been shown that changes in the expression pattern of specific genes are indicative of different pathological processes. It is also possible to distinguish between different types and stages of diseases, for example, various forms of cancer, heart disease, neuropsychiatric disorders, and the causes of infertility.[8,14] Another application in molecular medicine is pharmacogenomics, the analysis of gene expression to predict the response of a patient to treatment with specific drugs, thus enabling the choice of the most appropriate treatment for each individual patient.[8,15]

The use of gene-expression biomarkers for the detection of specific, external stimuli is a further field of application. Our group is interested in the misuse of drugs for growth-promoting purposes in human sports and animal husbandry. There are numerous reports about gene-expression changes caused by the use of anabolic substances in different tissues and species. In cattle, several promising candidate genes have been proposed for the detection of the misuse of anabolic substances; IGF-1, for example, has highly abundant expression in liver and muscle.[16–19] A further group of promising candidate genes are the receptors for specific substances, for example, the steroid hormone receptors or the β-adrenergic receptors in different tissues.[20–22] A lot of promising organs for biomarker discovery after hormone application have been reported: for example, uterus, ovary, prostate, vaginal epithelial cells, liver, muscle, and blood.[23–29] Most of these tissue samples have to be taken at the slaughterhouse or in the surgery room. In human sports, there is only a limited number of tissues available to trace the misuse of anabolic substances, for example, blood, urine, and hair. However, there are also reports regarding gene expression in human hair follicle cells and primate blood, suggesting that analysis of gene-expression changes caused by anabolic substances is feasible in humans.[27,29,30] All these examples demonstrate the potential of biomarker research at the mRNA level in different veterinarian and human research fields.

18.1.3 Biomarkers at the miRNA Level

miRNAs are small, regulatory RNA molecules that are involved in the regulation of mRNA expression and hence influence almost all physiological processes and metabolic pathways.[31] Dysregulation of miRNAs could be correlated with several different human pathologies, for example, diabetes, liver disease, or human cancer.[12,32–34] In this context, recent studies have revealed that specific miRNAs could be "the" upcoming biomarkers in clinical diagnostics. miRNAs show good suitability for biomarker research as they appear to be expressed in a developmental, disease, and tissue-specific manner, which is not the case for other established biomarkers. In contrast

to mRNAs, miRNAs are more stable[35] and less sensitive to RNAse exposure and, besides the transcriptional processing forms, from primary miRNA (pri-miRNA) to pre-miRNA to mature miRNA, no further modifications have been described. In contrast to the possibility of mRNA splice variants occurring from a single gene, for the mature miRNA no variants of the same molecule are known, thus facilitating accurate detection.[13,36] In cancer diagnostics, expression profiles of the so-called "oncomirs" (miRNAs, which are implicated in the formation of malignancies) have already proved their superiority over mRNA profiles.[37,38] For example, Lu and coworkers[12] were able to discriminate gastrointestinal cancer tissue from nongastrointestinal cancer tissue by characterization of specific miRNA profiles. This was not possible when screening the same biological samples for around 16,000 mRNAs. Besides diagnosis, screening of miRNA expression gives exceptional insights in to disease progression, for example, differentiation stages, developmental lineage of tumors, or response to therapy.[36,38]

miRNAs have also been discovered as circulating cell-free nucleic acids in the body fluids (e.g., blood, breast milk) of healthy and diseased individuals.[39–42] In addition, levels of circulating miRNAs in plasma have been linked to cancer (e.g., prostate, B-cell lymphoma) and other diseases (e.g., inflammatory bowel disease).[43–45] Mitchell et al.[43] showed that it is possible to identify prostate cancer patients by measuring the plasma levels of miR-141. In B-cell lymphoma, miR-21 was proposed as a promising biomarker because its serum abundance appears to be associated with patients' survival.[44] In blood, miRNAs are thought to be secreted from normal or tumor cells in microvesicles.[40] These findings could pose a breakthrough in the field of medical diagnostics as this would offer a possibility for prognostic information and early disease detection that is of minimal invasion.

Most studies that address circulating miRNAs as disease markers are targeting those that are originating from tumor cells as secretory products. However, these could also have physiological and regulative functions. For example, a notable number of miRNAs, which are known to play important functions in the immune system are found in breast milk (e.g., miR-155, miR-181a, miR-181b).[46,47] It is believed that miRNAs together with other immune-related agents contained in breast milk, like IgA and leucocytes, are responsible for the development of the immune system of the newborn baby by influencing the intestine.[41] Even though the underlying processes are not yet clarified, once more, the positive effect of breastfeeding for the health of the offspring can be supported by those observations.

18.2 miRNAs AS BIOMARKERS IN SURVEILLANCE OF ILLEGAL USE OF ANABOLIC STEROIDS

Even though the most common field of interest concerning miRNAs lies in human medical research, the concept of establishing biomarkers has also been introduced into veterinary medicine, for example, in the surveillance of illegal use of steroidal growth promoters.[19,24–27] Steroid hormones are known to alter gene expression and might also influence the expression of miRNAs. Recently, an innovative study investigated the effects of the anabolic combination of trenbolone acetate plus estradiol on miRNA abundance in bovine liver.[48] miRNA RT-qPCR arrays for gene-expression screening followed by statistical validation of results established an expression profile characterized by an upregulation of miR-29c, miR-130a, and miR-103 and a downregulation of miR-34a, miR-181c, miR-20a, and miR-15a. Using principal components analysis (PCA) as the biostatistical method of choice for pattern recognition (see Chapter 16), a separation on the basis of the miRNA expression profile between the untreated control group and treatment group could be shown. The significance of the group separation can be maximized, when integrating additionally significant mRNA expression results together with the miRNA.[48] It can be seen from these results, that the combination of gene expression results from mRNA and miRNA might be an upcoming integrative approach to use for the specific generation of gene-expression patterns as biomarkers for anabolic treatment screening.

18.3 TECHNICAL CHALLENGES IN miRNA BIOMARKER RESEARCH

RT-qPCR is the current gold standard for sensitive and reproducible miRNA gene-expression analysis. It is also established as the method of choice for validating results from holistic approaches, such as high-throughput sequencing (e.g., NGS), microarrays, or PCR array experiments.[36,49] However, the nature of miRNA molecules poses a challenge for reliable analytics.

The combination of short length of mature miRNAs (~22 nt) and a heterogeneous GC content poses a challenge for cDNA synthesis and primer and probe design since these results in significant difference in the melting temperatures of different miRNAs.

The sequence of interest is not only present in the mature miRNA, but also in the precursor sequences, the pri-miRNA and pre-miRNA.

The members of one miRNA family (e.g., let-7 family) usually differ by just one nucleotide, mostly at the 3′ end of the sequence.

There are no specific guidelines for miRNA expression data analysis and normalization.

Strategies to deal with those challenges have been published and are being intensively discussed (summarized in refs.[35,36,49–51]). Not only are the properties of the molecule challenging for established technical procedures, but also sample matrices pose additional problems. Especially in clinical research, patient samples (e.g., tumor samples) are frequently available as formalin-fixed paraffin-embedded (FFPE) tissues. From mRNA expression studies it is known that this type of tissue conservation is challenging because RNA from FFPE tissue is often cross-linked or degraded. Also, qPCR efficiency could be inhibited by formalin fixation.[52,53] Fortunately, various studies show that, in contrast to mRNAs, miRNAs seem to be less affected and more stable. Specialized extraction kits for FFPE samples could be used for miRNA expression analysis and reliable qPCR results could be gained.[53,54] Also, sample preparation from blood to perform RT-qPCR analysis of circulating miRNAs is challenging. Therefore, specified RNA isolation protocols and optional pre-amplification steps are required to deal with the low amounts of miRNAs present in plasma or serum. Additionally, inhibitors of qPCR present in blood (e.g., albumin) must be removed. An established method for analyzing cell-free miRNAs in plasma and serum is presented elsewhere.[42]

18.4 BIOINFORMATICS

As already mentioned, there is no single gene-expression biomarker for any given disorder or clinical situation. In most cases, multiple biomarkers must be present to distinguish between specific diseases, disease states, or treatments, hence a biomarker pattern consisting of various mRNA and miRNA transcripts, must be available. An important question is how to deal with these data to get the desired information. The best way seems to be the construction of clusters using methods for dimension reduction combined with pattern recognition technologies to visualize the gene-expression pattern in two- or three-dimensional graphs.[25,46] There are different multivariate analysis methods available, which are used for biomarker selection and validation, namely hierarchical clustering analysis (HCA) and PCA.

18.4.1 Hierarchical Cluster Analysis

The most popular method for the visualization of gene groups or treatment patterns is hierarchical clustering. An advantage of hierarchical clustering compared to the direct visualization methods is that a high dimensionality of the data set, represented by a large number of genes and samples, is reduced to a convenient two-dimensional representation of subject similarities.[55,56] HCA is the classification of similar objects into different groups, or more precisely, the partitioning of a data set into subsets, called clusters. The goal is to create clusters that share some common trait that is a matchable expression pattern. Hierarchical clustering can be performed either for the genes (comparing biological sample expression profiles) or for the biological samples (comparing

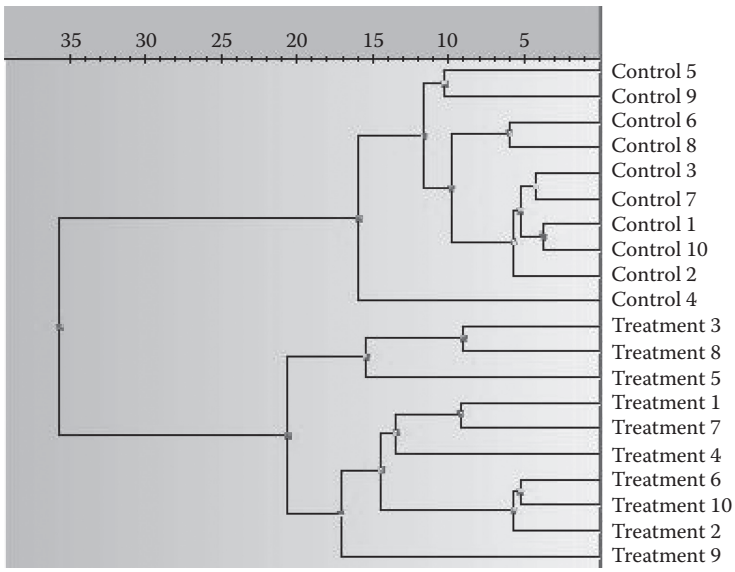

FIGURE 18.1 HCA of a set of 10 significantly regulated genes between 10 untreated control animals and 10 animals, treated with steroid hormones.

gene-expression profiles). HCA uses distance measure to identify pairs of animals showing high similarity based on the expression of different genes (Figure 18.1). Within many steps the animals with the highest similarity are merged in a cluster, and then the process is repeated. The result of the analysis is a tree dendrogram displaying the distances between the individuals based on the expression of genes.[56,57]

Using hierarchical clustering a tree dendrogram can either be designed for the measured genes (in all samples) or for the samples (based on all measured expressed genes). Using a heatmap analysis these two classifications can be combined, resulting in a two-dimensional color-coded description of the whole experimental matrix. It displays in a very convenient way all samples versus gene expression where each tile is colored with a different intensity according to all available data. Figure 18.2 shows a heatmap created from a set of 10 regulated genes in 20 animals (10 untreated control calves and 10 calves treated with steroid hormones). In both figures applying clustering methods, a clear separation of the two treatment groups underlying hormone-dependent physiological expression pattern changes upon selected biomarkers are visualized. In the two-dimensional heatmap, additionally the gene clusters with comparable regulation kinetics are obvious.

18.4.2 Principal Components Analysis

A further useful biostatistical and visualization method to group data is principal component analysis (PCA). PCA is a mathematical procedure that converts a multidimensional data set into a lower number of variables called principal components (PC).[57,58] The classification of the genes is based on unscaled Cq values and the overall changes of the gene-expression magnitudes.[59] The first principal component (PC1) represents the most significant PC, while gene-expression changes or variations in expression profiles are contained in the subsequent PCs. Inspecting the PC2, we see that treated and untreated individuals form two clusters that reflect the common biological functions and physiological processes of its members (Figure 18.3). Each analyzed animal will be represented by one spot which results from diminishing all significantly regulated genes of one specific sample to two PC. Variance from experimental study design conditions is

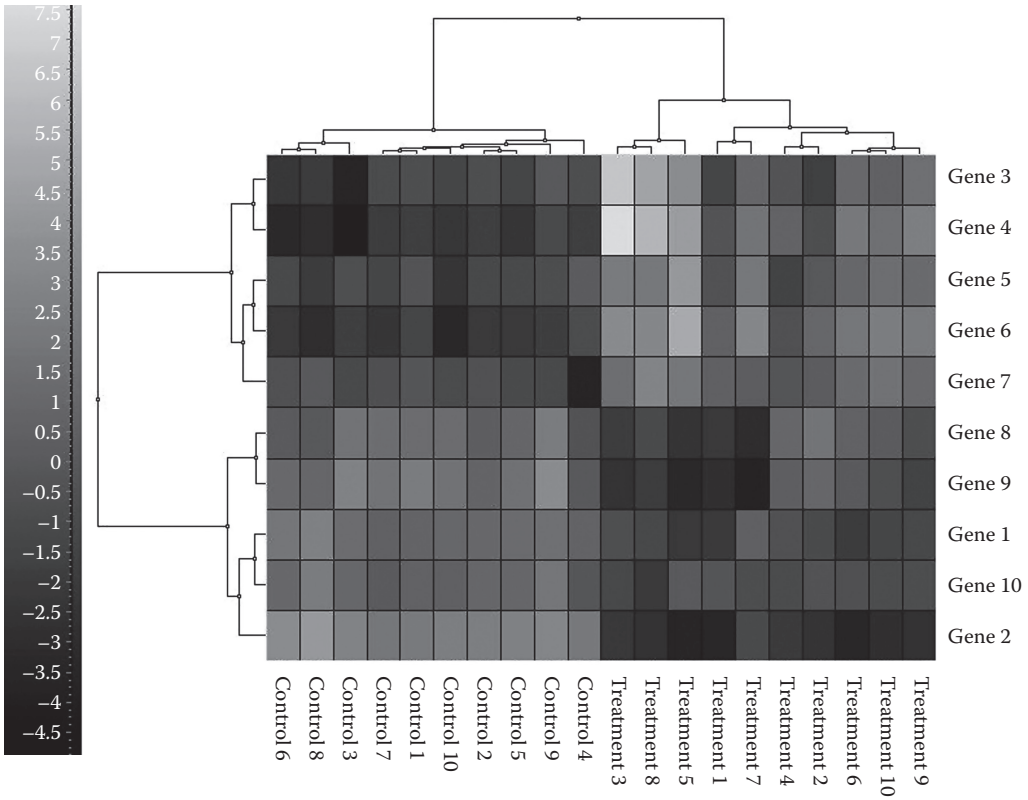

FIGURE 18.2 (**See color insert.**) Heatmap analysis of a set of 10 significantly regulated genes between 10 untreated control animals and 10 animals, treated with steroid hormones.

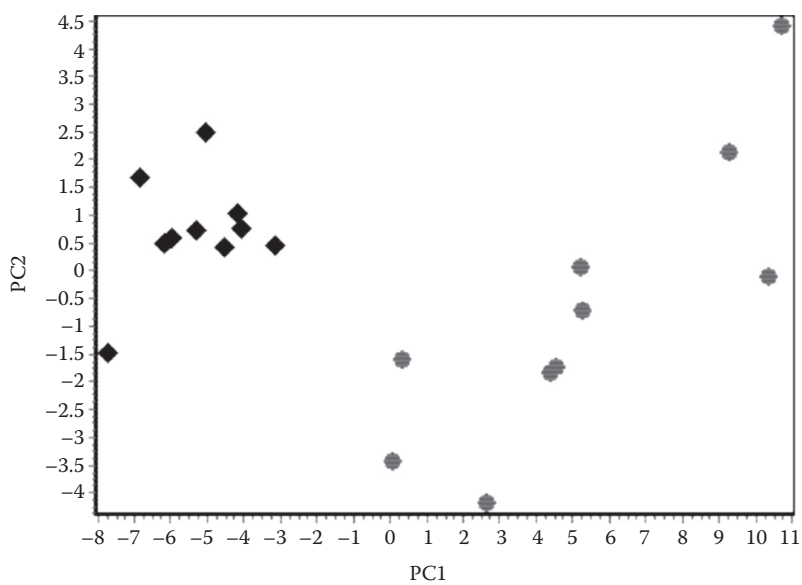

FIGURE 18.3 PCA of a set of 10 significantly regulated genes between 10 untreated control animals and 10 animals, treated with steroid hormones.

expected to be systematic, while confounding variance is expected to be random. Since the last PCs are derived from a very small amount of information, they can be considered to include noise or random information and can, therefore, be ignored. In this way, PCA can be a very efficient tool to separate systematic effects from noise.[56]

PCA has effectively been employed to visualize a treatment pattern in bovine tissues.[24–26,48,60] In bovine liver, PCA results obtained from mRNA and miRNA expression in combination showed a better separation between the groups than by employing the results from each individual transcript type. In Figure 18.3, a PCA of a set of 10 significantly regulated genes between 10 untreated control animals and 10 animals treated with steroid hormones is displayed. The clear separation of the two treatment groups indicate that PCA is a good tool for pattern recognition in gene-expression biomarker research.

The advantage of the PCA in comparison to the HCA methods is obvious. PCA allows a much clearer recognition and more precise differentiation of the treatment groups, because the commonalities in gene-expression pattern are visualized by the symbol interspaces in two dimensions.

18.5 SOFTWARE TOOLS

There are multiple software tools available to perform HCA, heatmaps, or PCA, either as standalone software or as packages freely available on the Internet.

The "Genex" software package offers a lot of tools to analyze mRNA and miRNA expression data in a correct and MIQE compliant way (according to Ref. [61]). Genex with its multiple functions helps to find and validate stable biomarkers (MultiD, Gothenburg, Sweden). More about the software, its functionality and the application of multidimensional data analysis is explained on the programmer's webpage (www.multid.com).

The "Genevestigator" software tool (www.genevestigator.com) aims at detecting specific patterns of expression in a multi-dimensional expression space by including a very large number of conditions processed from thousands of microarrays. The intuitive interface allows users easily to obtain lists of potential biomarker genes that can then be further validated using Genevestigator tools or in the laboratory. The classical clustering method (HCA) is for the grouping of genes according to their global pattern. Genevestigator provides several tools for clustering array data or meta-profiles.[6,7] The gene similarity is measured across all arrays or conditions. A dendrogram is applied to the clustered matrix and indicates relationships between clusters. More advanced biclustering is a method that identifies groups of genes that have similar profiles in a subset of conditions, irrespective of their profile similarity in the other conditions. Recent studies have shown that biclustering performs better than methods that require similarity over all conditions.[6,7]

A further method to discover and validate expressed biomarkers is to use R programming language, summarized in the "Bioconductor" project database (www.bioconductor.org). Bioconductor is an open source, open development software project to provide tools for the analysis and comprehension of high-throughput genomic data.[62] There are multiple packages and meta-data packages available, which provide the analysis of various data sources, for example, DNA, mRNA, miRNA, transcriptomics, microarray, real-time RT-PCR, sequence, or SNP. The broad goal is to provide widespread access to a full range of powerful statistical and graphical methods for the analysis of transcriptomics data. For real-time PCR data analysis and normalization a bundle of projects are available, for example, "HTqPCR," "qpcrNorm," "SLqPCR," or "ddCT" (summarized in "The qPCR library—Analysis of real-time PCR data using R"—http://www.dr-spiess.de/qpcR.html).[63] Further specialized packages for multidimensional expression analysis, PCA, HCA, or biomarker discovery are available in the database, for example, "BioMark" or "optBiomarker" project (http://www.bioconductor.org/help/search/index.html?q=biomarker).

Comparing the software packages, "Genex" and "Genevestigator" are the more user friendly, because they are working on a windows-based environment. The "Bioconductor" packages

expect an advanced operator who is able to handle and modify the text-based input script lines. Further, the graphical output of the results is limited in style and generally very simple in appearance.

18.6 CONCLUSION

We have described how biomarkers can be discovered from quantitative mRNA and miRNA transcript studies using RT-qPCR data obtained from various hormone treatment experiments in farm animals. The application of new transcriptomics technologies has resulted in the discovery of new, regulated transcripts and yielded potential biomarkers or biomarker patterns. But one critical point in biomarker discovery is the heterogeneity in the population and the variance of the biological samples itself. The application of integrative functional informatics as a novel direction of biomarker identification and validation seems to be very promising. Hence, the quantity of analyzed transcripts, on various levels and in multiple organs, in combination with the applied statistical method will have an impact on the informative value and the validity of the biomarkers.

Despite this enormous potential, so far none of the biomarker candidates described is included in veterinary screening or routine diagnostics. There is still a lack of validation of these discovered candidates in multiple organisms, in various environments, under changing conditions, and for veterinary research in multiple breeds. The existence of potential biomarkers is opening new insights in molecular diagnostics, an auspicious track to individualized treatment, or translated to human studies to future personalized medicine.

REFERENCES

1. Ilyin, SE, Belkowski, SM, Plata-Salaman, CR. 2004. Biomarker discovery and validation: Technologies and integrative approaches. *Trends Biotechnol.* 22:8, 411–6.
2. Hulka BS. Overview of biological markers. 1990. In: *Biological Markers in Epidemiology*. eds Hulka BS, Griffith JD, Wilcosky TC. Oxford University Press, Oxford, UK.
3. Biomarkers Definitions Working Group. 2001. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clin. Pharmacol. Ther.* 69:3, 89–95.
4. Mayeux, R. 2004. Biomarkers: Potential uses and limitations. *NeuroRx.* 1:2, 182–8.
5. Naylor, S. 2003. Biomarkers: Current perspectives and future prospects. *Expert. Rev. Mol. Diagn.* 3:5, 525–9
6. Hruz, T, Laule, O, Szabo, G, Wessendorp, F, Bleuler, S, Oertle, L, Widmayer, P, Gruissem, W, Zimmermann, P. 2008. Genevestigator v3: A reference expression database for the meta-analysis of transcriptomes. *Adv. Bioinformatics.* 2008, 420747.
7. Zimmermann, P, Laule, O, Schmitz, J, Hruz, T, Bleuler, S, Gruissem, W. 2008. Genevestigator transcriptome meta-analysis and biomarker search using rice and barley gene expression databases. *Mol. Plant* 1:5, 851–7.
8. Sandvik, AK, Alsberg, BK, Norsett, KG, Yadetie, F, Waldum, HL, Laegreid, A. 2006. Gene expression analysis and clinical diagnosis. *Clin. Chim. Acta* 363:1–2, 157–64.
9. Robosky, LC, Robertson, DG, Baker, JD, Rane, S, Reily, MD. 2002. *In vivo* toxicity screening programs using metabonomics. *Comb. Chem. High Throughput. Screen.* 5:8, 651–62.
10. Wang, Z, Gerstein, M, Snyder, M. 2009. RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10:1, 57–63.
11. Marioni, JC, Mason, CE, Mane, SM, Stephens, M, Gilad, Y. 2008. RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* 18:9, 1509–17.
12. Lu, J, Getz, G, Miska, EA, Alvarez-Saavedra, E, Lamb, J, Peck, D, Sweet-Cordero, A et al. 2005. MicroRNA expression profiles classify human cancers. *Nature* 435:7043, 834–8.
13. Wang, K, Zhang, S, Marzolf, B, Troisch, P, Brightman, A, Hu, Z, Hood, LE, Galas, DJ. 2009. Circulating microRNAs, potential biomarkers for drug-induced liver injury. *Proc. Natl. Acad. Sci. USA* 106:11, 4402–7.
14. Quackenbush, J. 2006. Microarray analysis and tumor classification. *N. Engl. J. Med.* 354:23, 2463–72.

15. Seo, D., Ginsburg, GS. 2005. Genomic medicine: Bringing biomarkers to clinical medicine. *Curr. Opin. Chem. Biol.* 9:4, 381–6.
16. Johnson, BJ, White, ME, Hathaway, MR, Christians, CJ, Dayton, WR. 1998. Effect of a combined trenbolone acetate and estradiol implant on steady-state IGF-I mRNA concentrations in the liver of wethers and the longissimus muscle of steers. *J. Anim. Sci.* 76:2, 491–7.
17. Pfaffl, MW, Daxenberger, A, Hageleit, M, Meyer, HH. 2002. Effects of synthetic progestagens on the mRNA expression of androgen receptor, progesterone receptor, oestrogen receptor alpha and beta, insulin-like growth factor-1 (IGF-1) and IGF-1 receptor in heifer tissues. *J. Vet. Med. A Physiol Pathol. Clin. Med.* 49:2, 57–64.
18. White, ME, Johnson, BJ, Hathaway, MR, Dayton, WR. 2003. Growth factor messenger RNA levels in muscle and liver of steroid-implanted and nonimplanted steers. *J. Anim. Sci.* 81:4, 965–72.
19. Reiter, M, Walf, VM, Christians, A, Pfaffl, MW, Meyer, HH. 2007. Modification of mRNA expression after treatment with anabolic agents and the usefulness for gene expression-biomarkers. *Anal. Chim. Acta* 586:1–2, 73–81.
20. Beermann DH. 2002. Beta-adrenergic receptor agonist modulation of skeletal muscle growth. *J. Anim. Sci.* 80:18–23.
21. Pfaffl, MW, Lange, IG, Meyer, HH. 2003. The gastrointestinal tract as target of steroid hormone action: Quantification of steroid receptor mRNA expression (AR, ERalpha, ERbeta and PR) in 10 bovine gastrointestinal tract compartments by kinetic RT-PCR. *J. Steroid Biochem. Mol. Biol.* 84:2–3, 159–66.
22. Sato, S, Nomura, S, Kawano, F, Tanihata, J, Tachiyashiki, K, Imaizumi, K. 2008. Effects of the beta2-agonist clenbuterol on beta1- and beta2-adrenoceptor mRNA expressions of rat skeletal and left ventricle muscles. *J. Pharmacol. Sci.* 107:4, 393–400.
23. Toffolatti, L, Rosa, GL, Patarnello, T, Romualdi, C, Merlanti, R, Montesissa, C, Poppi, L, Castagnaro, M, Bargelloni, L. 2006. Expression analysis of androgen-responsive genes in the prostate of veal calves treated with anabolic hormones. *Domest. Anim. Endocrinol.* 30:1, 38–55.
24. Riedmaier, I, Tichopad, A, Reiter, M, Pfaffl, MW, Meyer, HH. 2009. Identification of potential gene expression biomarkers for the surveillance of anabolic agents in bovine blood cells. *Anal. Chim. Acta* 638:1, 106–13.
25. Riedmaier, I, Becker, C, Pfaffl, MW, Meyer, HH. 2009. The use of omic technologies for biomarker development to trace functions of anabolic agents. *J. Chromatogr. A* 1216:46, 8192–9.
26. Riedmaier, I, Tichopad, A, Reiter, M, Pfaffl, MW, Meyer, HH. 2009. Influence of testosterone and a novel SARM on gene expression in whole blood of *Macaca fascicularis*. *J. Steroid Biochem. Mol. Biol.* 114:3–5, 167–73.
27. Riedmaier, I, Reiter, M, Tichopad, A, Pfaffl, MW, Meyer, HH. 2011. The potential of bovine vaginal smear for biomarker development to trace the misuse of anabolic agents. *Exp. Clin. Endocrinol. Diabetes* 119:2, 86–94.
28. Reiter, M, Pfaffl, MW, Schonfelder, M, Meyer, HH. 2008. Gene expression in hair follicle dermal papilla cells after treatment with stanozolol. *Biomark. Insights.* 4:1–8.
29. Reiter, M, Luederwald S, Pfaffl, MW, Meyer, HH. 2008. First steps towards a new screening method for anabolic adrogenic agents in human hair follicle. *Doping Journal* 5:3.
30. Reiter, M, Pfaffl, MW, Schoenfelder M, Meyer, HH. 2009. Gene expression in hair follicle dermal papilla cells after treatment with stanozolol. *Biomark. Insights.* 4:1–8.
31. Lee, RC, Feinbaum, RL, Ambros, V. 1993. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 75:5, 843–54.
32. Esquela-Kerscher, A., Slack, FJ. 2006. Oncomirs—MicroRNAs with a role in cancer. *Nat. Rev. Cancer* 6:4, 259–69.
33. Pandey, AK, Agarwal, P, Kaur, K, Datta, M. 2009. MicroRNAs in diabetes: Tiny players in big disease. *Cell Physiol. Biochem.* 23:4–6, 221–32.
34. Farazi, TA, Spitzer, JI, Morozov, P, Tuschl, T. 2011. miRNAs in human cancer. *J. Pathol.* 223:2, 102–15.
35. Becker, C, Hammerle-Fickinger, A, Riedmaier, I, Pfaffl, MW. 2010. mRNA and microRNA quality control for RT-qPCR analysis. *Methods* 50:4, 237–43.
36. Benes, V, Castoldi, M. 2010. Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available. *Methods* 50:4, 244–9.
37. Bartels, CL, Tsongalis, GJ. 2009. MicroRNAs: Novel biomarkers for human cancer. *Clin. Chem.* 55:4, 623–31.
38. Heneghan, HM, Miller, N, Kerin, MJ. 2010. MiRNAs as biomarkers and therapeutic targets in cancer. *Curr. Opin. Pharmacol.* 10:5, 543–50.

39. Laterza, OF, Lim, L, Garrett-Engele, PW, Vlasakova, K, Muniappa, N, Tanaka, WK, Johnson, JM et al. 2009. Plasma MicroRNAs as sensitive and specific biomarkers of tissue injury. *Clin. Chem.* 55:11, 1977–83.
40. Kosaka, N, Iguchi, H, Ochiya, T. 2010. Circulating microRNA in body fluid: A new potential biomarker for cancer diagnosis and prognosis. *Cancer Sci.* 101:10, 2087–92.
41. Kosaka, N, Izumi, H, Sekine, K, Ochiya, T. 2010. MicroRNA as a new immune-regulatory agent in breast milk. *Silence.* 1:1, 7.
42. Kroh, EM, Parkin, RK, Mitchell, PS, Tewari, M. 2010. Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). *Methods* 50:4, 298–301.
43. Mitchell, PS, Parkin, RK, Kroh, EM, Fritz, BR, Wyman, SK, Pogosova-Agadjanyan, EL, Peterson, A et al. 2008. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc. Natl. Acad. Sci. USA* 105:30, 10513–8.
44. Lawrie, CH, Gal, S, Dunlop, HM, Pushkaran, B, Liggins, AP, Pulford, K, Banham, AH et al. 2008. Detection of elevated levels of tumour-associated microRNAs in serum of patients with diffuse large B-cell lymphoma. *Br. J. Haematol.* 141:5, 672–5.
45. Ng, EK, Chong, WW, Jin, H, Lam, EK, Shin, VY, Yu, J, Poon, TC, Ng, SS, Sung, JJ. 2009. Differential expression of microRNAs in plasma of patients with colorectal cancer: A potential marker for colorectal cancer screening. *Gut* 58:10, 1375–81.
46. Thai, TH, Calado, DP, Casola, S, Ansel, KM, Xiao, C, Xue, Y, Murphy, A et al. 2007. Regulation of the germinal center response by microRNA-155. *Science* 316:5824, 604–8.
47. Bala, S, Marcos, M, Szabo, G. 2009. Emerging role of microRNAs in liver diseases. *World J. Gastroenterol.* 15:45, 5633–40.
48. Becker, C, Riedmaier, I, Reiter, M, Tichopad, A, Pfaffl, MW, Meyer, HH. 2011. Changes in the miRNA profile under the influence of anabolic steroids in bovine liver. *Analyst* 136:6, 1204–9.
49. Meyer, SU, Pfaffl, MW, Ulbrich, SE. 2010. Normalization strategies for microRNA profiling experiments: A "normal" way to a hidden layer of complexity? *Biotechnol. Lett.* 32:12, 1777–88.
50. Schmittgen, TD, Lee, EJ, Jiang, J, Sarkar, A, Yang, L, Elton, TS, Chen, C. 2008. Real-time PCR quantification of precursor and mature microRNA. *Methods* 44:1, 31–8.
51. Mestdagh, P, Van, VP, De, WA, Muth, D, Westermann, F, Speleman, F, Vandesompele, J. 2009. A novel and universal method for microRNA RT-qPCR data normalization. *Genome Biol.* 10:6, R64.
52. Masuda, N, Ohnishi, T, Kawamoto, S, Monden, M, Okubo, K. 1999. Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples. *Nucleic Acids Res.* 27:22, 4436–43.
53. Andreasen, D, Fog, JU, Biggs, W, Salomon, J, Dahslveen, IK, Baker, A, Mouritzen, P. 2010. Improved microRNA quantification in total RNA from clinical samples. *Methods* 50:4, S6–9.
54. Li, J, Smyth, P, Flavin, R, Cahill, S, Denning, K, Aherne, S, Guenther, SM, O'Leary, JJ, Sheils, O. 2007. Comparison of miRNA expression patterns using total RNA extracted from matched samples of formalin-fixed paraffin-embedded (FFPE) cells and snap frozen cells. *BMC. Biotechnol.* 7:36.
55. Lee, G, Rodriguez, C, Madabhushi, A. 2008. Investigating the efficacy of nonlinear dimensionality reduction schemes in classifying gene and protein expression studies. *IEEE/ACM. Trans. Comput. Biol. Bioinform.* 5:3, 368–84.
56. Bergkvist, A, Rusnakova, V, Sindelka, R, Garda, JM, Sjogreen, B, Lindh, D, Forootan, A, Kubista, M. 2010. Gene expression profiling—Clusters of possibilities. *Methods* 50:4, 323–35.
57. Beyene, J, Tritchler, D, Bull, SB, Cartier, KC, Jonasdottir, G, Kraja, AT, Li, N et al. 2007. Multivariate analysis of complex gene expression and clinical phenotypes with genetic marker data. *Genet. Epidemiol.* 31 Suppl 1:S103–9.
58. Kubista, M, Andrade, JM, Bengtsson, M, Forootan, A, Jonak, J, Lind, K, Sindelka, R et al. 2006. The real-time polymerase chain reaction. *Mol. Aspects Med.* 27:2–3, 95–125.
59. Vandesompele, J, Kubista, M, Pfaffl, MW. 2009. Reference gene validation software for improved normalization. In: *Real-Time PCR: Current Technology and Applications*. eds Logan J, Edwards K, Saundes N. Caister Academic Press, Norfolk, UK.
60. Becker, C, Riedmaier, I, Reiter, M, Tichopad, A, Groot, MJ, Stolker, AA, Pfaffl, MW, Nielen, MF, Meyer, HH. 2011. Influence of anabolic combinations of an androgen plus an estrogen on biochemical pathways in bovine uterine endometrium and ovary. *J. Steroid Biochem. Mol. Biol* 125(3):192–201.
61. Bustin, SA, Benes, V, Garson, JA, Hellemans, J, Huggett, J, Kubista, M, Mueller, R et al. 2009. The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55:4, 611–22.

62. Gentleman, RC, Carey, VJ, Bates, DM, Bolstad, B, Dettling, M, Dudoit, S, Ellis, B et al. 2004. Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol.* 5:10, R80.
63. Ritz, C, Spiess, AN. 2008. qpcR: An R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. *Bioinformatics* 24:13, 1549–51.