

STANDARD ARTICLE

Differences in miRNA differential expression in whole blood between horses with sarcoid regression and progression

Lucia Unger¹  | Vidhya Jagannathan² | Alicja Pacholewska^{1,2} | Tosso Leeb² | Vinzenz Gerber¹

¹Swiss Institute of Equine Medicine, Department of Clinical Veterinary Medicine, Vetsuisse Faculty, University of Bern, Bern, Switzerland

²Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, Switzerland

Correspondence

Lucia Unger, Swiss Institute of Equine Medicine, Vetsuisse Faculty, University of Bern and Agroscope, Länggassstrasse 124, 3012 Bern, Switzerland.
Email: lucia.unger@vetsuisse.unibe.ch

Funding information

Swiss Institute Research Funds, Grant/Award Number: ismequine.ch; Swiss Institute of Equine Medicine; University of Bern; Swiss Institute of Equine Medicine; University of Bern

Background: Currently no methods are available to predict the clinical outcome of individual horses with equine sarcoid (ES) disease.

Objective: To investigate if whole blood microRNA (miRNA) profiles can predict the long-term development of ES tumors.

Animals: Five horses with regression and 5 with progression of ES lesions monitored over 5-7 years and 5 control horses free of ES for at least 5 years.

Methods: For this cohort study, RNA extracted from whole blood samples from the regression, progression, and control groups was used for high throughput sequencing. Known and novel miRNAs were identified using miRDeep2 and differential expression analysis was carried out by the DESeq2 algorithm. Target gene and pathway prediction as well as enrichment and network analyses were conducted using TarBase, mirPath, and metaCore from GeneGo.

Results: Fourteen miRNAs were differentially expressed between regression and progression groups after accounting for the control condition: 4 miRNAs (28.6%) were upregulated and 10 miRNAs (71.4%) were downregulated with >2-fold change. Seven of the 10 downregulated miRNAs are encoded in an miRNA cluster on equine chromosome 24, homologous to the well-known 14q32 cluster in humans. Their target genes show enrichment for pathways involved in viral carcinogenesis.

Conclusions and Clinical Importance: Whole blood miRNA expression profiles are associated with long-term ES growth in horses and warrant further validation as prognostic biomarkers in a larger study cohort. Deregulation of miRNAs on equine chromosome 24 might represent a trigger for ES development.

KEYWORDS

equine cancer, equine sarcoid (ES) disease, high throughput sequencing, microRNA, prognostic biomarker

1 | INTRODUCTION

Equine sarcoid (ES) disease is the most common neoplastic skin condition of horses.^{1,2} Its pathogenesis is complex with host genetic susceptibility and bovine papilloma virus (BPV) 1 and 2 infection representing the most important inciting factors. The risk for being affected by ES

may be influenced by specific alleles or haplotypes at the major histocompatibility complex,³⁻⁵ but recent population genetics and genome-wide association studies suggest that susceptibility to ES has a polygenic basis.⁶⁻⁸ Once affected by ES, the clinical course of the disease can differ dramatically among individuals. Approximately 48% of single occult and verrucous sarcoids in young horses may show complete spontaneous regression over 5-7 years.⁹ Such milder forms of ES, however, can transform into aggressive types such as fibroblastic, mixed fibroblastic, or malevolent sarcoids. They may spread extensively or be

Abbreviations: BPV, bovine papilloma virus; ES, equine sarcoid; HPV, human papilloma virus; LRT, likelihood ratio test; miRNA, microRNA; PCA, principal component analysis; RIN, RNA integrity number.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2018 The Authors. *Journal of Veterinary Internal Medicine* published by Wiley Periodicals, Inc. on behalf of the American College of Veterinary Internal Medicine.

located in anatomically difficult areas, and thus compromise the health, welfare, and value of the affected animal.¹⁰

The diverse clinical forms of ES are still poorly understood. Although it is well accepted that exogenous factors such as trauma, including diagnostic biopsies or surgical intervention, can trigger or exacerbate the disease,^{11,12} far less is known about the endogenous factors that may influence disease outcome. Unlike other neoplastic diseases, histological characterization is not useful to discriminate between milder and more aggressive sarcoid phenotypes; increased density of dermal fibroblasts is common to all forms and other microscopic alterations are too variable to differentiate clinical types.¹³ The DNA of BPV is consistently present in ES, but BPV viral load and abnormal p53 expression do not appear to be associated with ES progression.¹⁴ In contrast, an *in vitro* study suggested that inhibition of BPV gene expression might reverse the transformed phenotype of sarcoid fibroblasts and therefore limit proliferation and invasiveness.¹⁵ Recently, abnormal DNA methylation patterns have been detected in sarcoid samples and these epigenetic changes might help to assess growth behavior of the tumors.¹⁶ In summary, it remains unknown why some horses show severe exacerbations and others are clinically stable or even experience spontaneous regression.

MicroRNAs (miRNAs) are small, 20-23 nucleotide long, noncoding RNA molecules that mainly downregulate the expression of their target genes. MicroRNA expression is dysregulated in various malignancies,¹⁷ and tumor-type specific miRNA fingerprints have been proposed as promising diagnostic and prognostic cancer biomarkers.¹⁸ Changes in miRNA expression have been observed in ES tissue and BPV-transformed equine fibroblast cell lines.^{19,20} Furthermore, miRNA fingerprints have prognostic potential in human papilloma virus (HPV)-associated cancers.²¹⁻²³

Thus, our aim was to compare miRNA expression profiles in whole blood of horses with regression versus progression of ES disease and ES-free control horses, assess their potential as noninvasive predictors of the course of ES disease, and predict their possible biological function.

2 | MATERIAL AND METHODS

2.1 | Study cohort

This study was fully approved by the Ethical Committee of the Canton of Bern (VD 2227.0, VD 2227.1 und VD 2227.2, BE110/15). Animals were client-owned (regression and progression groups) or clinic-owned horses (control group) that were cared for according to local regulations. Whole blood samples collected into EDTA tubes for a previous longitudinal study⁹ were available for discovery of prognostic miRNA fingerprints. They had been taken from 61 Franches-Montagnes horses that had been diagnosed with ES at the age of 3 years (1st examination and blood collection). After 5-7 years, a standardized clinical follow-up with a focus on ES lesion behavior was performed on all horses. Horses were defined as suffering from ES based on the characteristic clinical appearance of skin lesions and excluded when the diagnosis of ES was unclear because of atypical appearance

or location of the lesions. Animals were then grouped based on progression versus regression of lesions. For this study, 5 horses with marked regression (complete disappearance of lesions at 2nd examination) and 5 horses with marked progression (exacerbation at 2nd examination with more aggressive phenotype[s]) were selected. Whole blood samples also were collected from a control group consisting of 5 clinic-owned horses ≥ 12 years of age that were free of sarcoid lesions for at least 5 years and free of other signs of systemic disease at the time of examination. In control horses, whole blood collection into EDTA tubes was performed by the end of the 5-year observation period. Because of the high costs associated with next generation sequencing, a low number of study subjects (5 horses per group) were selected for this exploratory cohort study. Because no general consensus currently exists on what the sample size should be for miRNA biomarker studies, a previously recommended approach was used.²⁴

2.2 | RNA sample preparation and sequencing of multiplexed small RNA samples

Whole blood samples used in this study were stored for long term (9-11 years) for horses in the ES regression and progression groups and for short term (6 weeks) for control horses at -80°C before the analysis for this study. Samples were handled according to a previously established protocol for long-term stored equine whole blood.²⁵ The blood was gently thawed on ice and transferred to a PAXgene blood RNA tube 16 hours before the start of RNA extraction with a PAXgene Blood miRNA Kit (Qiagen, Hombrechtikon, Switzerland). The RNA concentrations were measured with Qubit RNA BR assay kit (ThermoFisher, Basel, Switzerland). The RNA integrity number (RIN) was assessed using Fragment Analyzer with the Standard Sensitivity RNA Analysis Kit (Advanced Analytical, Heidelberg, Germany). A total of 14.6-54.6 ng of small RNA was used for stranded single-end library preparation (NEBNext Multiplex Small RNA Library Prep Set for Illumina) and sequenced (HiSeq3000, Illumina).

2.3 | Raw data preprocessing

An average of approximately 17 million reads was collected for each of the 15 libraries. Raw sequencing data quality was assessed using FastQC software.²⁶ Truncated adapter sequences were reported by FastQC to be over-represented. Cutadapt (v 1.8)²⁷ was used to trim the full and truncated adapter sequences and to filter out reads < 15 bp with the following options: `--trim-n a<adapters> FASTA file> -m 15 aq 20`, in which “-m” denotes the minimum fragment length in bp and “-q” denotes the minimum read quality score encoded in `asci (phred quality + 33)` format. The raw sequence dataset is available at the European Bioinformatics Institute <http://www.ebi.ac.uk/ena/data/view/PRJEB27174>.

2.4 | MicroRNA annotation

Sequencing reads that passed quality control were mapped to the horse reference genome (EquCab2) available from NCBI using the miRDeep2 (v0.0.7) package.²⁸ The processed reads from all 15 FASTQ

files were mapped by the mapper.pl script of the miRDeep2 tool using a configuration file containing a list of all FASTQ file names and a 3-letter code for each sample (used to label each read according to the phenotype [regression, progression, and control]). The module mapped the reads to the genome with Bowtie1²⁹ keeping only the alignments with 0 mismatches (option -n) in the seed region and reads that did not map to >5 different loci in the genome (option -m). The module was run with the following parameters: -d -v -u -e -h -o 16 -m -n -q -p. This procedure generated nonredundant reads with the counts, source, and genomic location of the mapped reads. The miR-Deep2 then was used to annotate putative novel miRNAs as well as validate predicted miRNAs from miRBase³⁰ using (1) mapped reads from the mapper.pl output, (2) equine miRBase predicted miRNAs, and (3) human miRBase predicted miRNAs.²⁸ The putative novel miRNAs predicted by miRbase were assigned a score based on the 5p and 3p support and secondary structure that conformed to biogenesis miRNAs. Putative novel miRNAs predicted by miRDeep2 with a score of 4 were retained for further analysis because the signal over noise ratio for this cutoff was >10 at 14.3. Known equine and human miRNAs were obtained from the miRBase database, release 21.³⁰ Furthermore, overlapping novel miRNAs (by at most 1 nucleotide) were merged by BEDtools.³¹ The code used to implement this approach is outlined in Supporting Information File S1.

2.5 | Differential expression analysis of miRNA

The expression level quantification of novel (including merged ones) and known miRNA, predicted by miRDeep2, was performed by the quantifier.pl script applying the parameters -k -j. This script produced a raw counts file that then was used as input into the DESeq R/Bioconductor package for further statistical analysis.³² Differential expression tests in DESeq were conducted only for miRNA genes with read counts in ≥ 1 of the samples. For DESeq2, the "DESeq" function was used to estimate size factors, dispersions, and finally fit a model. The DESeq2 likelihood ratio test (LRT) was used to contrast the conditions (regression - control) - (progression - control) and obtain differential expressed miRNA among the 3 levels. The bioconductor package DegReport was used to visualize the results.³³ MiRNAs were considered as differentially expressed if the absolute value of log₂-fold change was >1 and the *P* value adjusted for multiple testing was <.05.

2.6 | MicroRNA target genes, enrichment, pathway, and network analysis

Enrichment analysis with GeneCodis³⁴⁻³⁶ was performed using target genes of differentially expressed miRNAs determined by TarBase v7.0 (Diana Tools),^{37,38} which catalogs only published, experimentally validated miRNA:gene interactions. Pathway prediction of the union of genes regulated by the differentially expressed miRNAs was performed by mirPath v.3 (Diana Tools).³⁸ MetaCore v.6.32 (Thomson Reuters) software was used to build networks of the differentially expressed miRNAs. Expanded networks (auto expand algorithm default: 50 nodes) were built using known and predicted targets of the differentially expressed miRNAs.

3 | RESULTS

3.1 | Study cohort

The regression group consisted of 3 mares and 2 stallions, and the progression group of 3 mares and 2 geldings. All horses were Franches-Montagnes horses and 3 years of age at the time of sampling. Sarcoid types consisted only of occult and verrucous lesions in both groups at the 1st presentation and verrucous, nodular, and fibroblastic lesions in the progression group at the 2nd examination. Anatomical sites and types of the lesions as well as treatment attempts are presented in Supporting Information Table S1. The control group consisted of 2 mares, 2 stallions, and 1 gelding aged 12-23 years (mean \pm SD, 16 \pm 4.36 years). The control group comprised 3 Franches-Montagnes and 2 Swiss Warmblood horses.

3.2 | RNA quality and quantity control

The concentration of RNA used for library preparation ranged from 14.0 to 59.6 ng/ μ L (mean \pm SD, 34.1 \pm 14.5 ng/ μ L). The RNA concentration in whole blood was 33.7 \pm 16.5 ng/ μ L in ES-affected horses and 35.1 \pm 11.2 ng/ μ L in control horses (Supporting Information Table S1). The RIN ranged from 1.0 to 8.3 (3.2 \pm 2.4). Mean RIN was 2.2 \pm 2.4 in whole blood of ES-affected horses and 5.2 \pm 0.6 in control horses (Supporting Information Table S1). Because the correlation between the coefficient of variation of expressed miRNAs and RIN quality of the input RNA was 0.2, control for RIN in differential gene expression analysis was not performed.

3.3 | Read sequence data quality control

Approximately 17 million reads from the 15 libraries were collected and after trimming for adapters and low-quality bases, approximately 11 million reads remained. The number of unique reads mapped by Bowtie1 to the horse reference genome was 981 012. A summary of miRNA read sequencing is shown in Supporting Information Table S2. Figure 1 displays the read length distribution for adapter and quality-trimmed reads for the 3 groups of samples.

3.4 | Read mapping and putative novel miRNA identification

The miRDeep2 algorithm reported 786 putative novel miRNAs with miR-Deep2 score between 0 and 10 (Supporting Information Table S3) and detected 271 novel miRNAs and 313 known miRNAs that passed the relatively stringent score cutoff of 4. These then were used for studying the differential expression with the DESeq2 tool. The novel miRNAs also were blasted against the Rfam database for potential hits to rRNA or tRNA, which did not result in any significant matches. These resulting high confidence novel miRNAs are listed in Supporting Information Table S4.

3.5 | Differential expression analysis with DESeq2

The 522 miRNAs with raw counts values quantified by miRDeep (from 15 samples) were used for differential analysis. After filtering rows with zero expression from the 522 miRNAs, 514 miRNA genes

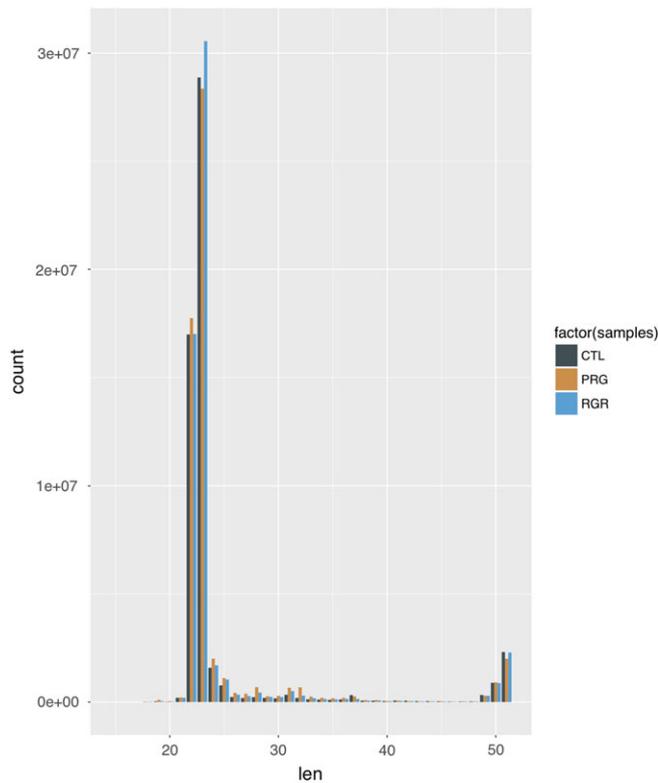


FIGURE 1 Read length distribution of all miRNA raw read libraries of control horses (CTL), horses with progressive sarcoids (PRG), and horses with regressive sarcoids (RGR). len, length

remained. A principal component analysis (PCA) plot was generated to determine the overall sample clustering based on the 200 most variable miRNAs. Three outliers were removed based on PCA clustering (Supporting Information Figures S1 and S2). The 3 outliers were ES_1595_E belonging to the control group, ES_1660_E belonging to the progression group, and ES_1677_E belonging to the regression group. The miRNA read counts from the remaining 12 samples were used for the differential expression.

As recommended by DESeq2 for this type of experimental design, a LRT was performed to test for differences in gene expression among

TABLE 1 Differentially expressed known and putative novel miRNAs in horses with regressive (regression-control) and progressive sarcoids (progression-control)

	baseMean	log2-FoldChange	P value	P adj
Eca-miR-381	63.6763	-3.4913	2.23E-09	9.97E-07
Eca-miR-134	39.5413	-3.5363	3.13E-07	6.99E-05
Eca-miR-127	138.0183	-3.6306	6.20E-07	9.24E-05
Eca-miR-382	15.4181	-4.2427	1.56E-05	.00174842
Eca-miR-379	10.6011	-5.0780	3.49E-05	.003
20_15133	37.1299	1.7902	.0001	.009
Eca-miR-107b	218.9891	-1.3709	.0002	.01
Eca-miR-146a	1252.0312	-1.7017	.0002	.01
Eca-miR-125a-5p	25.2910	1.1047	.0007	.03
Eca-miR-432	2.9624	-5.3878	.0007	.03
Eca-miR-24	1371.2551	0.0419	.0009	.03
Eca-miR-323-5p	14.6754	-3.2733	.001	.04
Eca-miR-1249	2192.4562	0.7270	.001	.04
Eca-miR-378	1234.2905	-1.5619	.001	.04

Abbreviation: P adj = adjusted P value.

the sample groups (in this case, a condition factor with 3 levels: regression, progression, and control). Fourteen differentially expressed miRNAs were detected between the regression and progression group which were accounted for by control condition (Table 1). The differential expression was identified with an adjusted P value <.05 and a fold change of approximately >2. Four miRNAs (28.6%) were upregulated and 10 miRNAs (71.4%) were downregulated in the regression group (Table 1). The degPlot function from DESeq2 was used to view the expression of the top 10 differentially expressed miRNAs (Figure 2). The degPatterns function from the R package DESeq2 was used to extract and plot genes that had a similar trend across the 3 conditions (control [CTL], progression [PRG], and regression [RGR] on the x-axis; Supporting Information Figure S3).

3.6 | MicroRNA target genes, enrichment, and pathway analysis

Using TarBase v7.0 (Diana Tools), 265 target genes with prediction score ≥ 0.9 (suggested threshold for stringent analysis) were predicted for the 14 differentially expressed miRNAs in the regression versus progression groups (Supporting Information Table S5). These target genes then were used for enrichment and pathway analysis. The results of the singular enrichment analysis of target genes using GeneCodis for gene ontology biological process and molecular function showed enrichment for cancer and cancer-related pathways (Supporting Information Table S6). MirPath v.3 (Diana Tools) predicted 67 pathways for all differentially expressed miRNAs in the regression versus progression group ($P < .05$) with the pathway “viral carcinogenesis” being between the top 2 pathways (Table 2).

3.7 | Biomarker-disease associations and network analysis for downregulated miRNAs in the regression versus progression group

To determine the comprehensive functional character of the differentially expressed miRNAs, ontology enrichment analysis was performed

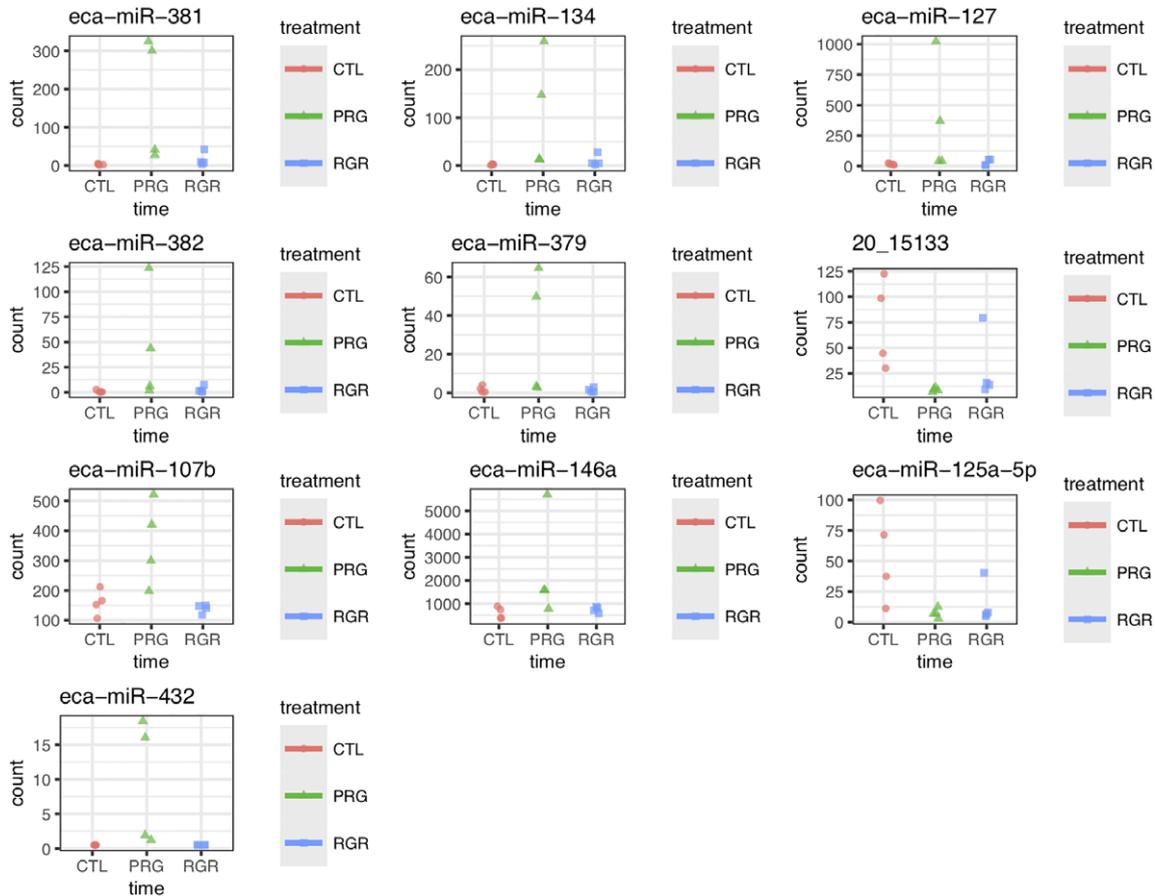


FIGURE 2 The top 10 differentially expressed miRNA patterns in control horses (CTL) and horses with progressive (PRG), and regressive sarcoids (RGR). The y-axis shows normalized counts of the respective miRNA

with target genes of up- and downregulated miRNAs by the MetaCore software algorithm of GeneGo. The direct target genes of downregulated miRNAs in the regression group were top sarcoma and connective/soft tissue neoplasm biomarkers according to GeneGo disease ontology (Figure 3).

The 10 miRNAs found to be downregulated in whole blood of the regression versus progression group are likely to be associated with a more favorable prognosis. Downregulation of some of their human

analogs in plasma also has been associated with a more favorable prognosis in human cancers (Supporting Information Table S7). Literature-based network analysis using GeneGo for downregulated miRNAs in whole blood of horses with regression of ES disease predicted miR-146a, miR-381, and miR-107 as central regulatory elements (Figure 4). MiR-146a, miR-381, and miR-107 deregulation also has been associated with HPV-associated cancers (Supporting Information Table S7). Furthermore, the expression of c-Myc, a major oncogene, is predicted to be regulated by 2 of the downregulated miRNAs (miR-134 and miR-382), which in human beings are described

TABLE 2 Pathway prediction for the differentially expressed miRNAs in horses with regressive sarcoids versus progressive sarcoids according to mirPath v.3 (Diana Tools) ($P \leq .05$). The number of miRNAs associated with each pathway and number of genes involved in each pathway are indicated. The top 10 pathways are listed

KEGG pathway	P value	Number of genes	Number of miRNAs
Proteoglycans in cancer	4.38E-18	106	10
Viral carcinogenesis	3.44E-09	103	10
Fatty acid metabolism	3.52E-09	22	4
Prion diseases	3.52E-09	15	7
Hippo signaling pathway	3.55E-07	70	9
Chronic myeloid leukemia	7.46E-07	45	8
Cell cycle	1.44E-06	72	9
Fatty acid biosynthesis	2.20E-06	6	4
Lysine degradation	2.20E-06	28	6
Hepatitis B	2.20E-06	70	9

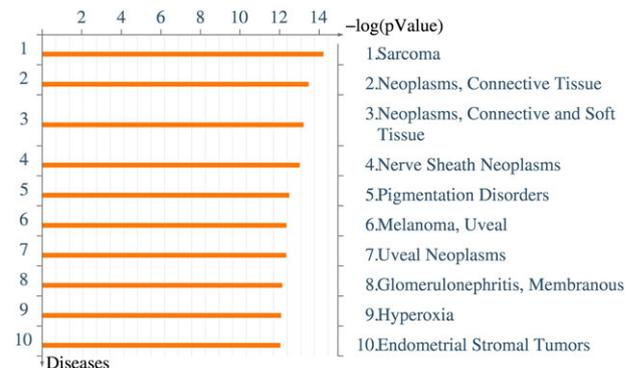


FIGURE 3 The top GeneGo disease ontology terms associated with direct targets of downregulated miRNAs in horses with regressive versus progressive sarcoids. The y-axis represents ranks according to the highest log of the P value in the x-axis

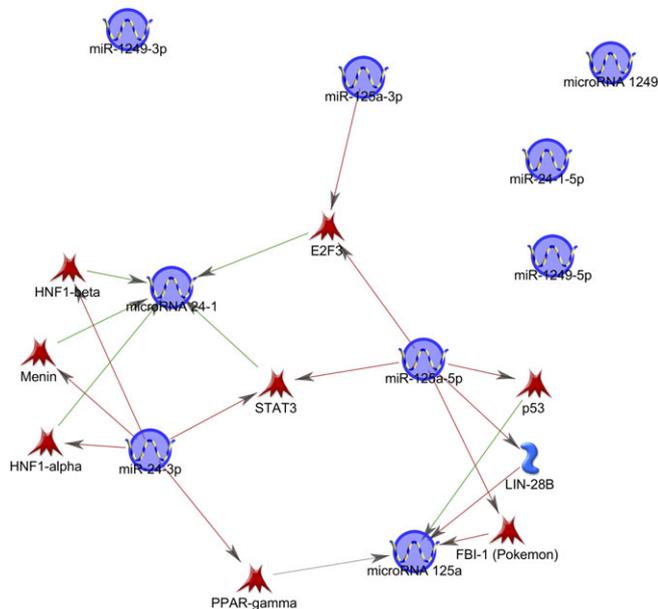


FIGURE 6 Network for upregulated miRNA genes in horses with regressive versus progressive sarcoids. According to network analysis using GeneGo miR-24 and miR-125a-5p are predicted to be major effector miRNAs in the group of upregulated miRNAs in the regression group by controlling expression of E2F3 and STAT3

cancer^{41–48} but can be upregulated in certain tumor types.^{49–51} The miRNA network on the 14q32 locus has been implicated in a variety of cancer types in humans and deregulations have been linked to clinical outcome in osteosarcoma,⁴⁷ ovarian cancer,⁴⁸ lung adenocarcinoma,⁴⁹ and melanoma.⁴¹ In the horse, the genomic organization of the miRNA genes corresponding to the human 14q32 locus is not perfectly conserved. However, 3 corresponding polycistronic miRNA clusters on equine chromosome 24 encode all of the 7 previously mentioned equine miRNAs.³⁹ In total, equine chromosome 24 harbors 40 miRNA genes in 4 polycistronic clusters.³⁹ Of the 56 miRNA genes in the human 14q32 cluster,⁴¹ 37 homologous equine miRNAs are found in the equine chromosome 24 region. Thus, we consider the miRNA cluster in the equine chr24: 42937011-42745108 region homologous to the human 14q32 miRNA cluster. Interestingly, human chromosome 14q32 is 1 of the most common HPV integration sites in cervical cancer. The corresponding miRNA cluster on equine chromosome 24 might represent a predilection site for BPV integration in ES disease.⁵² The differential expression of miRNAs belonging to the miRNA cluster on chromosome 24 in whole blood of horses with sarcoids also might reflect a dysregulation at the tumor tissue level: Eca-miR-381 and eca-miR-134 both are upregulated in sarcoid tissue of horses requiring treatment.²⁰ Ours is the first study to report an association of dysregulation of an miRNA cluster on chromosome 24 with cancer in horses.

Some of the differentially expressed miRNAs in our study have been linked to HPV-induced carcinogenesis in cancers of humans, as also is predicted by pathway analysis: miR-381 belongs to HPV core miRNAs and is downregulated in cervical squamous cell carcinoma and HPV-associated head and neck squamous cell carcinoma.^{53,54} In human oropharyngeal carcinoma, HPV positivity is associated with miR-107 downregulation, possibly because of HPV-enhanced immune

response, whereas overexpression of miR-107 was associated with decreased overall and disease-free survival.²³ The main HPV oncoproteins, E6 and E7, can strongly influence miRNA expression in HPV-induced cancers: miR-24 is upregulated by E6 and E7 expression and may promote cell proliferation by targeting the cell cycle inhibitor p27.⁵⁵ In contrast, miR-1249 is downregulated in HPV-positive human keratinocytes in response to HPV16 E6/E7 oncoprotein expression.⁵⁶ It needs to be further evaluated whether these mechanisms of miRNA dysregulation also apply to BPV oncoproteins in ES disease. Recently, miRNA deregulation has been attributed to the presence of BPV genomes in BPV-transformed equine fibroblasts.¹⁹

According to network analysis, eca-miR-24 and eca-miR-125a-5p turned out to be central regulatory elements, presumably by regulating the expression of E2F3 and STAT3, and their upregulation has been associated with advantageous outcome of ES disease. Both miRNAs already have been shown to be dysregulated at tissue or cell culture level: eca-miR-24 has been found to be downregulated in BPV-transformed equine fibroblasts.¹⁹ Eca-miR-125a-5p is highly expressed in sarcoid tissue and downregulated in serum of horses with aggressive, fibroblastic ES lesions (unpublished data—manuscript in preparation).

In our study, miRNA fingerprints in horses with ES regression versus progression were evaluated in whole blood. Typically, serum is the preferred body fluid for noninvasive miRNA profiling.⁵⁷ The value, advantages, and limitations of whole blood samples as input material for miRNA studies are controversial, because changes in blood cell counts can strongly influence miRNA expression (such as to an immune response against neoplastic disease).⁵⁸ Other studies, however, highlight the potential of whole blood compared to serum: whole blood contains higher miRNA concentrations⁵⁹ and important sources of pre-analytical errors related to serum processing, such as hemolysis, can be avoided.⁶⁰ Furthermore, tumor-specific exosomal miRNA might only be detectable in serum if a well-perfused neoplastic mass of substantial size is present, whereas cellular blood cell-derived miRNA fingerprints may already be detectable in sufficient amounts at earlier stages of tumor development,⁶¹ and with masses that are poorly perfused, such as ES lesions. Moreover, sarcoids in horses are not accompanied by changes in blood cell counts unless marked superficial ulceration is present, which was not the case in any of the initially mild ES lesions in our study. Thus, the use of whole blood as input material for miRNA analysis seemed superior to serum, particularly in light of the presence of only small and mild ES lesions in the regression and progression groups at the time of sample collection.

As for any retrospective study, our study has some limitations. Diagnosis of ES in our study was clinical, even though a definitive diagnosis of ES requires histopathology. However, clinical diagnosis has good reliability and because trauma of any nature carries the risk of exacerbating ES,^{10,12} biopsies were not performed. Furthermore, surgical excision of the relatively mild lesions in our study (Supporting Information Table S1) was not indicated in any of the horses at the time of first presentation.

Because of the retrospective and longitudinal nature of our study, blood samples of different quality collected at different time points throughout the study period had to be utilized. Long-term stored whole blood samples collected at the beginning of the study period

were used in the ES-affected horses, in contrast to the short-term stored samples from control horses collected by the end of the study period. This difference in sample quality explains the higher RIN values in the control group. Still, we considered having a reliable control group consisting of daily monitored clinic-owned horses, which did not develop sarcoids at any point throughout the 5-year study period, more important than including a control group from the previously described Franches-Montagnes study population,⁹ from which a blood sample collected at the beginning of the study period would have been available, which was however only examined on 2 occasions throughout the 5-7 year study period. In the latter group, smaller transiently occurring sarcoid lesions might have been easily missed and could have substantially biased the results of the study. Comparison of miRNA differential expression in blood of horses with ES collected at the beginning, at several time points during, and at the end of the study period might have shed further light on changes of miRNA fingerprints during the course of ES disease and might have allowed us to draw additional conclusions regarding their possible biological functions. Further longitudinal studies therefore should investigate miRNA differential expression repeatedly and closely during the clinical course of ES.

Some reports state that miRNA quantification is influenced by overall RNA integrity,⁶² and that RNA degradation compromises the reliability of miRNA analysis.⁶³ More recent studies, however, show that miRNA profiles themselves are not affected by the extent of RNA degradation.^{64,65} In our study, we could confirm that RIN did not have any impact on differential gene expression analysis. Thus, our study shows that despite the compromised quality of a subset of blood samples because of long-term storage, a robust miRNA study still could be conducted. This observation encourages the use of valuable blood samples from bioarchives for future longitudinal studies.

The mean age of the control horses was significantly higher than the mean age of horses in the ES groups. Sarcoids often develop in young horses,⁴ thus only ES-free middle-aged and old horses (≥ 12 years) were included in the control group. We did not want to include younger age-matched controls because of the risk of developing ES lesions and becoming false-negative controls. Age may affect miRNA expression profiles in human beings.^{66,67} Its effect on the results of the present study was considered to be less important than the need for a reliable control group.

Our study provides a set of 14 blood-derived miRNAs significantly associated with prognosis of ES disease in a limited number of samples with markedly favorable (regression) versus aggressive (progression) clinical courses. These miRNA fingerprints warrant further validation as prognostic biomarkers in a larger study cohort. Admittedly, to date, miRNA differential expression analysis remains an elaborate diagnostic technique and for now may not serve as a practical test and prognostic biomarker in equine practice. However, in the future, qRT-PCR may replace next generation sequencing to assess miRNA expression levels, which would render the procedure less cost prohibitive and time consuming. Additional studies focusing on correlation of whole blood miRNA expression and miRNA and target gene expression at the tissue level as well as monitoring of changes of miRNA expression over time may shed more light on the speculated biological functions of these blood cell-derived miRNAs, particularly, if dysregulation of the miRNA cluster on equine chromosome 24 by

integration of the BPV genome is a possible mechanism that drives disease progression. Regardless of the biological function of blood-derived miRNAs, these potential biomarkers might offer novel perspectives in the clinical setting, such as for prepurchase examinations or therapeutic decision-making.

ACKNOWLEDGMENTS

The authors thank Alessandra Ramseyer, Dominik Burger, and Christoph Koch for their support and assistance with sample and animal data collection. The authors also thank Muriel Fragnière and Sabrina Schenk for expert technical assistance. The Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit of the University of Bern are acknowledged for performing the whole genome resequencing experiments and providing high-performance computing infrastructure. The work was done at Swiss Institute of Equine Medicine, University of Bern. The presented study was funded by the Swiss Institute of Equine Medicine Research funds (ismequine.ch). Presented as an abstract at the 2017 European College of Equine Internal Medicine Congress, Budapest, Hungary.

CONFLICT OF INTEREST DECLARATION

Authors declare no conflict of interest.

OFF-LABEL ANTIMICROBIAL DECLARATION

Authors declare no off-label use of antimicrobials.

INSTITUTIONAL ANIMAL CARE AND USE COMMITTEE (IACUC) OR OTHER APPROVAL DECLARATION

This study was fully approved by the Ethical Committee of the Canton of Bern (VD 2227.0, VD 2227.1 und VD 2227.2, BE110/15).

HUMAN ETHICS APPROVAL DECLARATION

Authors declare human ethics approval was not needed for this study.

ORCID

Lucia Unger  <https://orcid.org/0000-0001-8238-8907>

REFERENCES

- Schaffer PA, Wobeser B, Martin LE, et al. Cutaneous neoplastic lesions of equids in the Central United States and Canada: 3,351 biopsy specimens from 3,272 equids (2000-2010). *J Am Vet Med Assoc.* 2013;242:99-104.
- Valentine BA. Survey of equine cutaneous neoplasia in the Pacific Northwest. *J Vet Diagn Invest.* 2006;18:123-126.
- Lazary S, Marti E, Szalai G, Gaillard C, Gerber H. Studies on the frequency and associations of equine leucocyte antigens in sarcoid and summer dermatitis. *Anim Genet.* 1994;25(Suppl 1):75-80.
- Brostrom H. Equine sarcoids. A clinical and epidemiological study in relation to equine leucocyte antigens (ELA). *Acta Vet Scand.* 1995;36:223-236.
- Brostrom H, Fahlbrink E, Dubath ML, et al. Association between equine leucocyte antigens (ELA) and equine sarcoid tumors in the

- population of Swedish halfbreds and some of their families. *Vet Immunol Immunopathol.* 1988;19:215-223.
6. Christen G, Gerber V, Dolf G, Burger D, Koch C. Inheritance of equine sarcoid disease in Franches-Montagnes horses. *Vet J.* 2014;199:68-71.
 7. Staiger EA, Tseng CT, Miller D, et al. Host genetic influence on papillomavirus-induced tumors in the horse. *Int J Cancer.* 2016;139:784-792.
 8. Jandova V, Klukowska-Rotzler J, Dolf G, et al. Whole genome scan identifies several chromosomal regions linked to equine sarcoids. *Schweiz Arch Tierheilkd.* 2012;154:19-25.
 9. Berruex F, Gerber V, Wohlfender FD, Burger D, Koch C. Clinical course of sarcoids in 61 Franches-Montagnes horses over a 5-7 year period. *Vet Q.* 2016;36:189-196.
 10. Knottenbelt DC. A suggested clinical classification for the equine sarcoid. *Pferdeheilkunde.* 2006;22:479-480.
 11. McConaghy FF, Davis RE, Reppas GP, et al. Management of equine sarcoids: 1975-93. *N Z Vet J.* 1994;42:180-184.
 12. Taylor S, Haldorson G. A review of equine sarcoid. *Equine Vet Educ.* 2013;25:210-216.
 13. Martens A, De Moor A, Demeulemeester J, et al. Histopathological characteristics of five clinical types of equine sarcoid. *Res Vet Sci.* 2000;69:295-300.
 14. Bogaert L, Van Poucke M, De Baere C, et al. Bovine papillomavirus load and mRNA expression, cell proliferation and p53 expression in four clinical types of equine sarcoid. *J Gen Virol.* 2007;88:2155-2161.
 15. Yuan Z, Gault EA, Campo MS, Nasir L. Different contribution of bovine papillomavirus type 1 oncoproteins to the transformation of equine fibroblasts. *J Gen Virol.* 2011;92:773-783.
 16. Semik E, Zabek T, Gurgul A, et al. Comparative analysis of DNA methylation patterns of equine sarcoid and healthy skin samples. *Vet Comp Oncol.* 2017;16:37-46.
 17. Croce CM. Causes and consequences of microRNA dysregulation in cancer. *Nat Rev Genet.* 2009;10:704-714.
 18. Iorio MV, Croce CM. MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. *EMBO Mol Med.* 2012;4:143-159.
 19. Terron-Canedo N, Weir W, Nicolson L, et al. Differential expression of microRNAs in bovine papillomavirus type 1 transformed equine cells. *Vet Comp Oncol.* 2016;15:764-774.
 20. Pawlina K, Gurgul A, Szmatala T, et al. Comprehensive characteristics of microRNA expression profile of equine sarcoids. *Biochimie.* 2017;137:20-28.
 21. Gao G, Gay HA, Chernock RD, et al. A microRNA expression signature for the prognosis of oropharyngeal squamous cell carcinoma. *Cancer.* 2013;119:72-80.
 22. Wong N, Khwaja SS, Baker CM, et al. Prognostic microRNA signatures derived from The Cancer Genome Atlas for head and neck squamous cell carcinomas. *Cancer Med.* 2016;5:1619-1628.
 23. Hui AB, Lin A, Xu W, et al. Potentially prognostic miRNAs in HPV-associated oropharyngeal carcinoma. *Clin Cancer Res.* 2013;19:2154-2162.
 24. Kok MGM, de Ronde MWJ, Moerland PD, Ruijter JM, Creemers EE, Pinto-Sietsma SJ. Small sample sizes in high-throughput miRNA screens: a common pitfall for the identification of miRNA biomarkers. *Biomol Detect Quantif.* 2018;15:1-5.
 25. Unger L, Fouche N, Leeb T, et al. Optimized methods for extracting circulating small RNAs from long-term stored equine samples. *Acta Vet Scand.* 2016;58:44.
 26. Andrews S. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. 2010. Accessed May 1, 2018.
 27. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal.* 2011;17:10-12. <http://journal.embnet.org/index.php/embnetjournal/article/view/200>.
 28. Friedlander MR, Mackowiak SD, Li N, et al. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* 2012;40:37-52.
 29. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10:R25.
 30. Griffiths-Jones S, Grocock RJ, van Dongen S, et al. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* 2006;34:140-144.
 31. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010;26:841-842.
 32. Anders S, McCarthy DJ, Chen YS, et al. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat Protoc.* 2013;8:1765-1786.
 33. Pantano L. DESeq: Report of DEG analysis. R package version 1.16.0. <https://bioconductor.org/packages/release/bioc/html/DESeq.html>: Bioconductor; 2017.
 34. Tabas-Madrid D, Nogales-Cadenas R, Pascual-Montano A. GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Res.* 2012;40:478-483.
 35. Nogales-Cadenas R, Carmona-Saez P, Vazquez M, et al. GeneCodis: interpreting gene lists through enrichment analysis and integration of diverse biological information. *Nucleic Acids Res.* 2009;37:317-322.
 36. Carmona-Saez P, Chagoyen M, Tirado F, Carazo JM, Pascual-Montano A. GENECODIS: a web-based tool for finding significant concurrent annotations in gene lists. *Genome Biol.* 2007;8:R3.
 37. Vlachos IS, Paraskevopoulou MD, Karagkouni D, et al. DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. *Nucleic Acids Res.* 2015;43:153-159.
 38. Vlachos IS, Zagganas K, Paraskevopoulou MD, et al. DIANA-miRPath v3.0: deciphering microRNA function with experimental support. *Nucleic Acids Res.* 2015;43:460-466.
 39. Kim MC, Lee SW, Ryu DY, Cui FJ, Bhak J, Kim Y. Identification and characterization of microRNAs in normal equine tissues by next generation sequencing. *PLoS One.* 2014;9:e93662.
 40. Yue P, Turkson J. Targeting STAT3 in cancer: how successful are we? *Expert Opin Investig Drugs.* 2009;18:45-56.
 41. Zehavi L, Avraham R, Barzilai A, et al. Silencing of a large microRNA cluster on human chromosome 14q32 in melanoma: biological effects of mir-376a and mir-376c on insulin growth factor 1 receptor. *Mol Cancer.* 2012;11:44.
 42. Lavon I, Zrihan D, Granit A, et al. Gliomas display a microRNA expression profile reminiscent of neural precursor cells. *Neuro Oncol.* 2010;12:422-433.
 43. Thyanithy V, Sarver AL, Kartha RV, et al. Perturbation of 14q32 miRNAs-cMYC gene network in osteosarcoma. *Bone.* 2012;50:171-181.
 44. Haller F, von Heydebreck A, Zhang JD, et al. Localization- and mutation-dependent microRNA (miRNA) expression signatures in gastrointestinal stromal tumours (GISTs), with a cluster of co-expressed miRNAs located at 14q32.31. *J Pathol.* 2010;220:71-86.
 45. Saito Y, Liang G, Egger G, et al. Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell.* 2006;9:435-443.
 46. Kelly AD, Haibe-Kains B, Janeway KA, et al. MicroRNA paraffin-based studies in osteosarcoma reveal reproducible independent prognostic profiles at 14q32. *Genome Med.* 2013;5:2.
 47. Sarver AL, Thyanithy V, Scott MC, et al. MicroRNAs at the human 14q32 locus have prognostic significance in osteosarcoma. *Orphanet J Rare Dis.* 2013;8:7.
 48. Zhang L, Volinia S, Bonome T, et al. Genomic and epigenetic alterations deregulate microRNA expression in human epithelial ovarian cancer. *Proc Natl Acad Sci U S A.* 2008;105:7004-7009.
 49. Nadal E, Zhong JJ, Lin JL, et al. A microRNA cluster at 14q32 drives aggressive lung adenocarcinoma. *Clin Cancer Res.* 2014;20:3107-3117.
 50. Luk JM, Burchard J, Zhang CS, et al. DLK1-DIO3 genomic imprinted microRNA cluster at 14q32.2 defines a stemlike subtype of hepatocellular carcinoma associated with poor survival. *J Biol Chem.* 2011;286:30706-30713.
 51. Costa FF, Bischof JM, Vanin EF, et al. Identification of microRNAs as potential prognostic markers in ependymoma. *PLoS One.* 2011;6:e25114.
 52. Wentzensen N, Vinokurova S, von Knebel Doeberitz M. Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. *Cancer Res.* 2004;64:3878-3884.

53. Lajer CB, Garnaes E, Friis-Hansen L, et al. The role of miRNAs in human papilloma virus (HPV)-associated cancers: bridging between HPV-related head and neck cancer and cervical cancer (vol 106, pg 1526, 2012). *Brit J Cancer*. 2017;117:1-2.
54. Lajer CB, Nielsen FC, Friis-Hansen L, et al. Different miRNA signatures of oral and pharyngeal squamous cell carcinomas: a prospective translational study. *Br J Cancer*. 2011;104:830-840.
55. McKenna DJ, Patel D, McCance DJ. miR-24 and miR-205 expression is dependent on HPV onco-protein expression in keratinocytes. *Virology*. 2014;448:210-216.
56. Harden ME, Prasad N, Griffiths A, Munger K. Modulation of microRNA-mRNA target pairs by human papillomavirus 16 oncoproteins. *MBio*. 2017;8:e02170-16.
57. Chen X, Ba Y, Ma L, et al. Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res*. 2008;18:997-1006.
58. Pritchard CC, Kroh E, Wood B, et al. Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. *Cancer Prev Res (Phila)*. 2012;5:492-497.
59. Patnaik SK, Yendamuri S, Kannisto E, Kucharczuk JC, Singhal S, Vachani A. MicroRNA expression profiles of whole blood in lung adenocarcinoma. *PLoS One*. 2012;7:e46045.
60. Kirschner MB, Kao SC, Edelman JJ, et al. Haemolysis during sample preparation alters microRNA content of plasma. *PLoS One*. 2011;6:e24145.
61. Hausler SFM, Keller A, Chandran PA, et al. Whole blood-derived miRNA profiles as potential new tools for ovarian cancer screening. *Brit J Cancer*. 2010;103:693-700.
62. Becker C, Hammerle-Fickinger A, Riedmaier I, Pfaffl MW. mRNA and microRNA quality control for RT-qPCR analysis. *Methods*. 2010;50:237-243.
63. Ibberson D, Benes V, Muckenthaler MU, Castoldi M. RNA degradation compromises the reliability of microRNA expression profiling. *BMC Biotechnol*. 2009;9:102.
64. Jung M, Schaefer A, Steiner I, et al. Robust microRNA stability in degraded RNA preparations from human tissue and cell samples. *Clin Chem*. 2010;56:998-1006.
65. Hall JS, Taylor J, Valentine HR, et al. Enhanced stability of microRNA expression facilitates classification of FFPE tumour samples exhibiting near total mRNA degradation. *Br J Cancer*. 2012;107:684-694.
66. Meder B, Backes C, Haas J, et al. Influence of the confounding factors age and sex on microRNA profiles from peripheral blood. *Clin Chem*. 2014;60:1200-1208.
67. Ameling S, Kacprowski T, Chilukoti RK, et al. Associations of circulating plasma microRNAs with age, body mass index and sex in a population-based study. *BMC Med Genomics*. 2015;8:61.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Unger L, Jagannathan V, Pacholewska A, Leeb T, Gerber V. Differences in miRNA differential expression in whole blood between horses with sarcoïd regression and progression. *J Vet Intern Med*. 2018;1-10. <https://doi.org/10.1111/jvim.15375>