Published online 25 February 2016

Nucleic Acids Research, 2016, Vol. 44, No. 8 ***3865–3877***
doi: 10.1093/nar/gkw116

Distribution of miRNA expression across human tissues

Nicole Ludwig[1], Petra Leidinger[1], Kurt Becker[2], Christina Backes[3], Tobias Fehlmann[3], Christian Pallasch[4,5], Steffi Rheinheimer[1], Benjamin Meder[6,7,8], Cord Stähler[9], Eckart Meese[1] and Andreas Keller[3,*]

[1]Institute of Human Genetics, Saarland University, Medical School, Homburg, Germany, [2]Institute of Anatomy and Cell Biology, Saarland University, Medical School, Homburg, Germany, [3]Chair for Clinical Bioinformatics, Saarland University, Saarbruecken, Germany, [4]Department I of Internal Medicine and Center of Integrated Oncology, University Hospital of Cologne, Cologne, Germany, [5]Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), Cologne, Germany, [6]Department of Internal Medicine III, University Hospital Heidelberg , 69120 Heidelberg, Germany, [7]German Center for Cardiovascular Research (DZHK), 69120 Heidelberg, Germany, [8]Klaus Tschira Institute for Integrative Computational Cardiology, D-69118 Heidelberg, Germany and [9]Siemens Healthcare, Hartmannstrasse 16, 91052 Erlangen, Germany

Received May 06, 2015; Revised February 17, 2016; Accepted February 17, 2016

ABSTRACT

We present a human miRNA tissue atlas by determining the abundance of 1997 miRNAs in 61 tissue biopsies of different organs from two individuals collected post-mortem. One thousand three hundred sixty-four miRNAs were discovered in at least one tissue, 143 were present in each tissue. To define the distribution of miRNAs, we utilized a tissue specificity index (TSI). The majority of miRNAs (82.9%) fell in a middle TSI range i.e. were neither specific for single tissues (TSI > 0.85) nor housekeeping miRNAs (TSI < 0.5). Nonetheless, we observed many different miRNAs and miRNA families that were predominantly expressed in certain tissues. Clustering of miRNA abundances revealed that tissues like several areas of the brain clustered together. Considering -3p and -5p mature forms we observed miR-150 with different tissue specificity. Analysis of additional lung and prostate biopsies indicated that inter-organism variability was significantly lower than inter-organ variability. Tissue-specific differences between the miRNA patterns appeared not to be significantly altered by storage as shown for heart and lung tissue. MiRNAs TSI values of human tissues were significantly ($P = 10^{-8}$) correlated with those of rats; miRNAs that were highly abundant in certain human tissues were likewise abundant in according rat tissues. We implemented a web-based repository enabling scientists to access and browse the data (https://ccb-web.cs.uni-saarland.de/tissueatlas).

INTRODUCTION

Knowing the expression and distribution of different molecule classes in tissues is essential for the understanding of both physiological and pathological mechanisms. The gene expression atlas (1), hosted at the European Bioinformatics Institute, collects gene expression patterns under different biological conditions in various organisms. Likewise, the Human Protein Atlas presents information on proteomes in various tissues (2). For the class of small non-coding nucleic acids, the so-called microRNAs or miRNAs, there is a lack of up-to-date databases showing their tissue-specific distribution. The first and as of now most comprehensive analysis of miRNA abundance in different tissues has been reported by Landgraf et al. in 2007 (3). This sequencing-based study reported 340 miRNAs in 26 organs. We recently investigated the miRNA repertoire of different blood cell types (4), already indicating a complex miRNA repertoire strongly dependent on the considered cell types. To improve the understanding of the miRNA abundance in human tissues, we now profiled 1997 different mature miRNAs for 61 tissues. In contrast to the previous catalogue of miRNAs in human tissues, we measured all miRNA profiles from only two different individuals to minimize inter-individual variability. We selected an array-based analysis to have a robust platform for determining the miRNA abundance. The applied Agilent microarray technology has been proven sensitive and, more important, reproducible in a recent comprehensive platform comparison (5). Using this technology, we achieved technical Pearson correlation co-

*To whom correspondence should be addressed. Tel: +49 681 302 68611; Fax: +49 681 302 58094; Email: andreas.keller@ccb.uni-saarland.de

efficients of between 0.97 and 1 for technical replicates in previous studies.

Here, we first characterize technical stability of our approach before we describe variations in the abundance of the miRNAs across tissues. To provide easy access to the tissue atlas, we implemented a web-based repository that also links results to important miRNA resources. This web service is freely available online at https://ccb-web.cs.uni-saarland.de/tissueatlas.

MATERIALS AND METHODS

Tissues and RNA extraction

Tissues analysed in this study originated from two male bodies. Both cadavers were obtained as anatomical gift to be dissected in a study of medicine under German law. The first body was from a 65-year-old male patient, who suffered from multiple myeloma, a cancer that forms in a type of white blood cells (plasma cells). The body was stored at 4°C upon arrival at the anatomical institute and tissue samples were collected the following day, i.e. 2 days post-mortem. In total, we analysed 24 different tissues, i.e. adipocytes, arachnoid mater, artery, colon, small intestine (ileum), dura mater, brain, urinary bladder, skin, myocardium, bone (rib), liver, lung, stomach, spleen, muscle, gall bladder, muscle fascia, epididymis, intercostal nerve, kidney, thyroid, testis and tunica albuginea of testis.

The second body was from a 59-year-old male individual, who died a natural death. The body was frozen at −20°C after arrival at the anatomical institute and dissected after 3 weeks of storage. Autopsy showed no signs of cancer. As we aimed at increasing the resolution our tissue atlas, we collected 37 samples including several sub-areas for different organs, i.e. nine brain areas (grey matter, white matter, frontal, temporal, occipital, nucleus caudatus, thalamus, pituitary gland and cerebellum), dura mater, spinal cord, nerve, artery, vein, myocard, muscle, lymph node, thyroid, esophagus, stomach, pancreas, duodenum, jejunum, colon, liver, three kidney areas (kidney unspecified, medulla and cortex), spleen, adrenal gland, prostate, testis, skin, adipocyte, lung, pleura and bone marrow.

To assess the influence of RNA degradation originating from different storage times of the tissue on the miRNA profile, we used normal lung and normal heart tissue that was stored in physiological salt solution at 4°C for 1, 2, 3, 7 and 14 days, before RNA isolation. To understand short-term effects on the miRNA pattern in a comprehensive manner, we analysed lung tissue from another individual. The following 16 time points were profiled: 0, 0.5, 1, 1.5, 2, 3, 4, 5, 6, 9, 12, 24, 36, 48, 72 and 96 h.

To estimate inter-individual variations, we exemplarily performed in-depth analysis for lung tissues. For 16 normal tissue biopsies from different individuals, the miRNA expression intensity was determined as for the two bodies and the samples from the degradation analysis.

RNA isolation and integrity

RNA was isolated using the miRNeasy Mini Kit (Qiagen) and the Qiagen tissue lyser using 7 mm stainless steel beads. Tissue samples were disrupted for 5 min 30 Hz (1800 oscillations/min) in Qiazol lysis reagent. Further purification was done according to manufacturer's instructions. Concentration and purity was measured using NanoDrop 2000 (Thermo Scientific). RNA integrity was measured using Bioanalyzer RNA Nano Chip (Agilent). As expected for autopsy samples, the RNA integrity values (RIN) ranged between 1.8 and 2.7.

miRNA profiling

Microarray analysis was performed using *SurePrint* 8 × 60K Human V19 *miRNA* microarrays (Agilent) that contain 2007 miRNAs of miRBase V19 (http://www.mirbase.org/), according to the manufacturer's instructions for the first corpse. For the second corpse, the most recent miRBAse v21 has been used and the analysis has been carried out on 1997 human miRNAs present in both versions. In brief, a total of 100 ng RNAs were processed using the miRNA Complete Labeling and Hyb Kit to generate fluorescently labelled miRNA. Microarrays were scanned with the Agilent Microarray Scanner at 3 μm in double path mode. Microarray scan data were further processed using Agilent Feature Extraction software. The raw expression intensity values are available for download at https://ccb-web.cs.uni-saarland.de/tissueatlas. Since the normalization may have an impact on the results, we performed all analyses on the raw data, normalized data by quantile normalization and by variance stabilizing normalization (6). For training the Variance Stabilized Normalization (VSN) model all samples and all miRNAs were used. The detailed results for the variance stabilizing normalization are provided in the supplementary material. To account for negative values (i.e. miRNAs that are not expressed, that may get a negative value due to background subtraction) a pseudo-count has been added. All calculations have been carried out in R version 3.0.2.

Tissue specificity index

To evaluate the variability of expression patterns, we calculated a tissue specificity index (TSI) for each miRNA analogously to the TSI 'tau' for mRNAs originally developed by Yanai et al. (7). This specificity index is a quantitative, graded scalar measure for the specificity of expression of a miRNA with respect to different organs. The values range from 0 to 1, with scores close to 0 represent miRNAs expressed in many or all tissues (i.e. housekeepers) and scores close to 1 miRNAs expressed in only one specific tissue (i.e. tissue-specific miRNAs). Specifically, the TSI for a miRNA j is calculated as

$$\mathrm{tsi}_j = \frac{\sum_{i=1}^{N}(1 - x_{j,i})}{N-1},$$

where N corresponds to the total number of tissues measured and $x_{j,\mathrm{i}}$ is the expression intensity of tissue i normalized by the maximal expression of any tissue for miRNA j.

Hierarchical clustering of tissues

To estimate the proximity of profiles from different tissues, hierarchical clustering analysis has been applied. To ac-

count for the high dynamic range of miRNAs, clustering has been performed on log expression intensities and miRNAs that are close to the background were removed. To extend the cluster analysis, the 100 most variable miRNAs have been selected. In each case, complete linkage hierarchical clustering using the Euclidian distance has been performed.

Expression of miRNA families

For estimating the tissue specificity of miRNA families, we extracted all miRNA families from the most recent miRBase version 21. For each miRNA precursor all mature forms have been considered as family members, duplicated mature miRNAs (e.g. coming from different precursors in the same family) have been counted once in order to minimize a potential bias introduced by multiple precursors. For discovering co-expressed miRNAs, Spearman correlation of intensity values between all pairs of miRNAs has been calculated. Network visualization has been performed in Cytoscape.

Conservation of tissue specificity

To compare conserved tissue specificity in humans and rats, we downloaded data from the Gene Expression Omnibus (GEO) series GSE52754, containing expression profiles for 55 different rat tissues that have been measured using Agilent microarrays (8). To match miRNAs we extracted all rat miRNA identifiers from the respective manuscript and matched them via a 100% sequence match. For matching miRNAs and matching tissues, we calculated and correlated the tissue specificity indices. To minimize artefacts introduced by normalization, we carried out all analyses on raw data. Since this analysis only addresses the question whether a miRNA is rather specific or a housekeeping miRNA, we also correlated the human and rat expression profiles using Spearman correlation.

Additional data from literature

In addition to the 44 tissue samples from the degradation and reproducibility analysis, the 16 individual lung cancer tissues and the 61 tissues from two bodies newly measured for this study, we searched the literature for other studies where normal tissues have been profiled. In the GEO (9), we found 1178 series related to miRNAs. Of these, 722 were from *Homo sapiens*. Excluding series with low sample count (below 20 samples), 302 series remained. After excluding studies from body fluids such as serum, plasma, blood or urine, we examined the remaining hits for availability of unaffected tissue measurements. The respective data tables were downloaded from GEO and all IDs were matched from the respective platform identifiers to miRBase Version 21 IDs. For the respective studies, raw and normalized data (VSN and quantile normalized) were added to our tissue atlas web repository. These include 43 samples from 9 tissues and 463 miRNAs from GSE11879, 40 samples measured for 709 miRNAs from normal gastric tissue from GSE23739, 48 benign prostate tissues measured for 480 miRNAs from series GSE54516 and 32 benign prostate tissues measured for 825 miRNAs from series GSE76260. The data have been used partially in the present manuscript, all data are included in the web-based tissue atlas resource.

RESULTS

In this work, we present the draft of a human tissue miRNome atlas. In the first part of the manuscript, we describe pre-analytics, investigating the general reproducibility of the miRNA profile measurements and also the effect of storage of tissues on miRNA profiles. In the pre-analytics consideration, we measured 44 tissue miRNomes. It is essential to understand respective variability to understand the biological variability of different tissue miRNomes.

In the second part, we describe the screening of all mature miRNAs from miRBase version 21 across different organs of two male bodies. We investigated miRNA expression in 24 different tissues from the first body and in 37 different tissues from the second body. To determine the miRNAs abundance in the different tissues, we utilized a TSI score, known from transcriptomics (7). Furthermore, we investigated the proximity of organs based on miRNA abundances by hierarchical clustering and co-expression analysis. To estimate inter-individual variations, we measured 16 additional miRNomes from control lung tissues and extracted further data sets from the GEO.

To provide researchers access to the first version of the miRNA tissue atlas, we implemented a web-based repository that is freely available at www.ccb.uni-saarland.de/tissueatlas.

Reproducibility of miRNA patterns

An important factor for estimating the biological variability is to understand the technical variability of the underlying profiling platform. Previously, we compared technical reproducibility of the two common platforms, microarrays (Agilent) and NGS (Illumina HiSeq) (10). In these experiments, we discovered an increased variability of miRNAs dependent of the sequencing library preparation. Similarly, we observed a strong bias based on the nucleotide composition of miRNAs (11). Of 10 replicated Agilent microarray measurements of the same individual, we calculated 10 * 9/2 pair-wise correlations of technical replicates. Minimal correlation was 0.998 and mean/median correlation 0.999, highlighting the high degree of technical reproducibility of the array platform. To translate these results on our tissue atlas and determine technical reproducibility of the array analysis, technical duplicates from nine randomly selected tissue samples from the second body were measured. The duplicates were processed at different days and have been measured on different arrays, each. Hierarchical cluster analysis shows that the technical replicates always clustered together showing that the applied technology was suited to provide reproducible results (Figure 1 shows the heat map for quantile-normalized data, Supplementary Figure S1 for VSN-normalized data). Altogether, we found high correlations between these technical replicates with the overall lowest correlations at 0.986 and 0.994 observed for liver tissue and pleura, respectively. Highest correlation of 0.999 was reached for the brain samples.

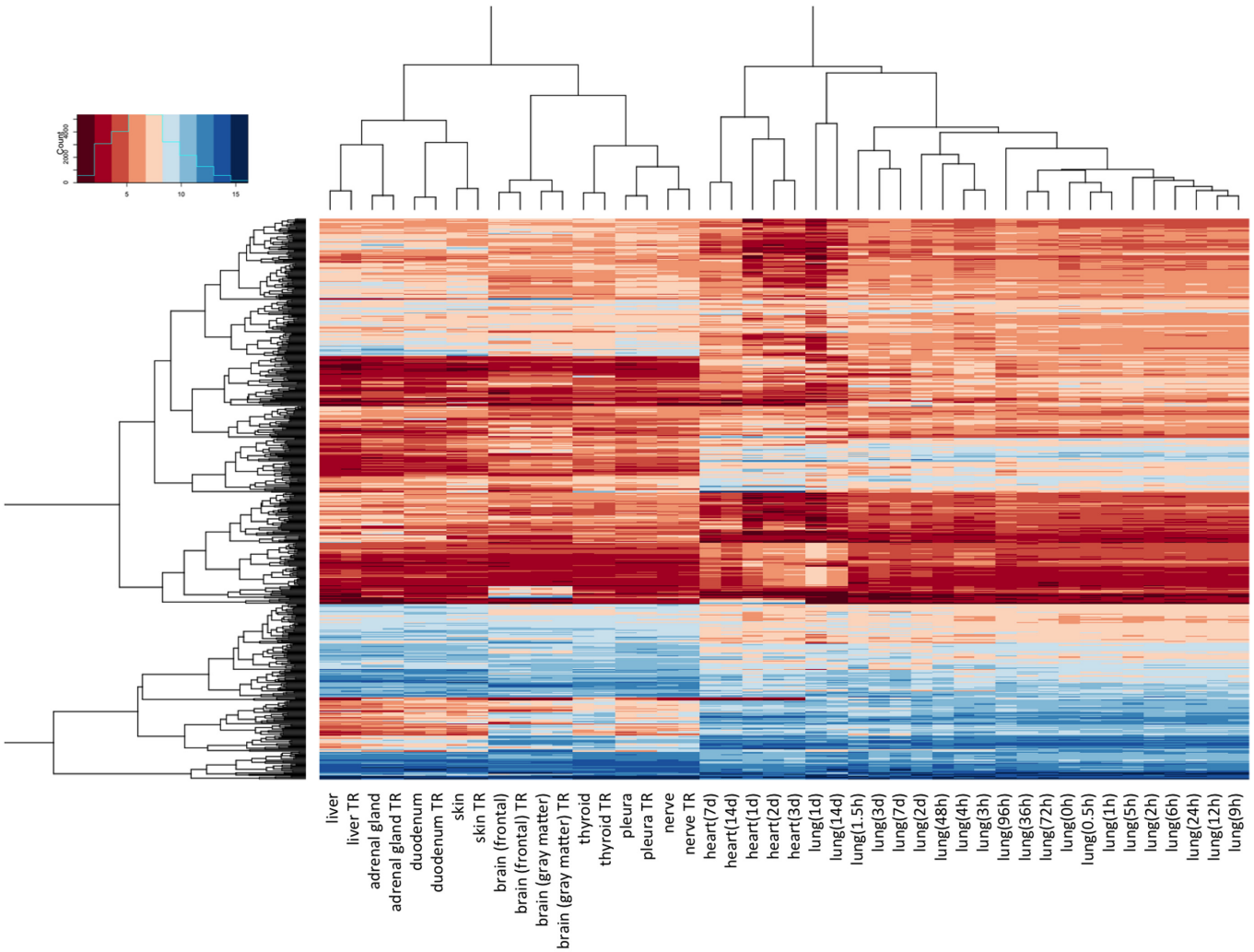

Figure 1. Hierarchical clustering of the 44 samples included in the stability and reproducibility study. Quantile normalized and $\log_2$ transformed expression intensity values were used for clustering. The intensity values and distribution are presented in the upper left corner. In the present heat map, heart and lung tissues cluster together on the right-hand side. Technical replicates (marked by 'TR' in the labels below the heat map) of other organs cluster together in each case in the left-hand side. For VSN-normalized data the same representation is provided in Supplementary Figure S1.

Stability of miRNA patterns in tissues

Measuring tissues of corpses the storage time prior to RNA extraction and a potential degradation of RNA may have an influence on the profiles. We exemplarily investigated the process for heart and lung tissue. Biopsies were taken from two individuals and have been stored for 1, 2, 3, 7 and 14 days at 4°C. Hierarchical cluster analysis shows that all lung and all heart samples each cluster together (Figure 1; Supplementary Figure S1). The duration of the storage was, however, not reflected in the clustering pattern indicating that a storage time between 1 and 14 days at 4°C has a limited influence on the overall miRNA tissue pattern.

We also performed the analysis with more dense time intervals within the first 3 days to understand short-term effects. For a lung tissue from a third individual 16 time points between 0 and 96 h were profiled. These biopsies clustered well together with the lung tissues from the second individual with storage time over 14 days. Again, no time curse could be recognized in the clustering pattern.

Remarkably, the results presented above describe the overall miRNA patterns. For single miRNAs still differences dependent on the storage could be observed. Thus, we calculated the TSI for all lung tissues and for all tissues in the pre-analytical study. With respect to lung tissues, large TSI values mean in this case not tissue specific but rather specific in one of the replicated measurements. We thus expect that TSI values of miRNAs from the lung tissue are low. Especially for five miRNAs we, however, calculated TSI values that are increased in lung tissue by at least 20%: hsa-miR-8069, hsa-miR-6821–5p, hsa-miR-4800–5p, hsa-miR-6775–5p, hsa-miR-5001–5p. For all miRNAs, TSI values from the pre-analytical step are summarized in Supplementary Table S1.

Frequency of miRNAs per tissue and tissue specificity of miRNAs

For each miRNA in each tissue, we determined its presence and frequency using the so-called present calls as determined by Agilent Feature Extraction software. Out of the 1997 different mature miRNAs, 633 (31.7%) were not detected in any of the tested tissues by the applied microarray technology. Out of the remaining 1364 miRNAs, 143 (10.5%) were found in all tissues. To present more comprehensive information on the tissue distribution of miRNAs, we utilized the miRNA TSI analogously to the mRNA TSI 'tau' that has successfully been employed by Yanai et al. (7). This index has a range of 0–1 with the score of 0 corresponding to ubiquitously expressed miRNAs (i.e. 'housekeepers') and a score of 1 for miRNAs that are expressed in a single tissue (i.e. 'tissue-specific' miRNAs). We calculated TSI for the 1364 miRNAs that have been detected in at least one tissue sample. For each miRNA, we compared TSI for the two bodies, for raw, quantile- and VSN-normalized data (Supplementary Table S2). Using the quantile-normalized data for the first body, 83.7% of all miRNAs showed an average abundance throughout the tissues with intermediate TSI values ranging from 0.15 to 0.85 (Figure 2A, Supplementary Figure S2A for VSN-normalized data). Only one miRNA (miR-3960) was ubiquitously expressed with a TSI < 0.15 and 222 miRNAs showed a highly tissue-specific expression with TSI > 0.85. For the second body, 88.8% of all miRNAs showed intermediate TSI values; one miRNA (miR-6089) showed a TSI < 0.15 and 152 miRNAs a TSI > 0.85 (Figure 2B, Supplementary Figure S2B for VSN-normalized data). The correlation of the VSN-normalized TSI values with the quantile-normalized TSI values was 0.88 ($P < 10^{-10}$).

The overall most tissue-specific miR-1–3p is presented in Figure 3. For all 61 samples raw-, quantile- and VSN-normalized expression intensities are presented as bar plot. Respective bar plots for all miRNAs can be generated using the online repository.

Clustering of tissue patterns and analysis of miRNA families

Beyond the analysis of single miRNAs, we determined the overall similarity/dissimilarity of the miRNA pattern between the different tissues. We performed hierarchical clustering of miRNAs and tissues using normalized expression intensities. We found two major clusters, the first of which containing mainly nervous system tissues and muscle tissues from both bodies. In the second cluster, the organs of the two individuals frequently did not cluster together (Figure 4A). Since the large number of miRNAs used for this clustering likely caused substantial noise, we restricted the clustering analysis to the 100 miRNAs with the highest data variance (Figure 4B). Here, we found three main clusters with the first one containing kidney, liver, stomach and small intestine of both bodies. The second cluster exclusively contained all brain tissue samples of both bodies and nervous system related tissue, i.e. spinal cord and dura mater. The third cluster contained thyroid, nerve, muscle, myocardium and colon each of both bodies. Other organs were found in different clusters, e.g. the lung samples and the brain coverings dura mater and arachnoid mater. For VSN-normalized data we observed a similar pattern, however, we found a stronger tendency of clustering of individuals in the different sub-clusters (Supplementary Figure S3).

To gain further insights into expression of tissue-specific miRNAs, we performed clustering with the 25 miRNAs displaying a TSI > 0.85 for both bodies in raw-, quantile- and VSN-normalized data (Figure 5). We found several groups of miRNAs with tissue-specific expression. In detail, we detected high expression of miR-133b, miR-133a-3p, miR-1–3p and miR-206 in both muscle samples and, with the exception of miR-206 also in both myocardial samples. Additionally, we found a cluster of four miRNAs specifically expressed in various brain tissues, i.e. miR-338–3p, miR-219a-5p, miR-124–3p and miR-9–5p. Another group of miRNAs, miR-507, miR-514a-3p and miR-509–5p was almost exclusively expressed in the testis samples. Besides these miRNA clusters, we also found single miRNAs that were expressed in a highly tissue-specific manner, i.e. miR-122–5p, miR-7–5p and miR-205–5p were each exclusively expressed in liver, pituitary gland and skin, respectively.

Tissue specificity of miRNA families

To further determine to what extend miRNA families show similar abundances in specific organs, we calculated the TSI not only for single miRNAs but also for mature miRNAs inside each miRNA family. Out of 187 miRNA families from the miRBase with at least two family members, we analysed 25 miRNA families with at least five mature forms (Figure 6A; Supplementary Table S3). We found several miRNA families with high TSI values including the above-mentioned mir-378 family with most of the family members showing a high abundance in muscle tissues and the myocardium. Similarly, the mir-506 family with 18 family members showed generally a high abundance in testis tissue while they were less expressed in other tissues. Other families, such as the mir-449 family with five members, did not show a common pattern in the different tissues: MiR-449c-3p was expressed specifically in spleen tissue, miR-449c-5p and -449b-5p in kidney and small intestine, miR-449a in lung, kidney and brain and miR-449b-3p in spleen. To extend this analysis we searched for miRNAs co-expression patterns in specific tissues. We used a high correlation cut-off and considered only miRNA-pairs with Pearson correlation exceeding 0.95. Altogether, we identified 73 miRNA pairs with tissue co-expression. In addition to pair-wise interactions, we also found sub-networks with at least four participants. The networks have been visualized using Cyto-Scape (Figure 6B). While we frequently observed co-expression among mature members of specific families (e.g. the mir-548 family), we also found correlations of miRNAs from different miRNA families. For example, miR-4312 was co-expressed with miRNAs from the let-7 family. Performing the same analysis with raw data, we detected an increased number of co-expressions, but generally confirmed the observation that has been based on the normalized data.

Tissue specificity of -3p and -5p mature forms

We asked whether -3p and -5p mature forms of miRNAs have different tissue specificity. To limit the bias of miR-

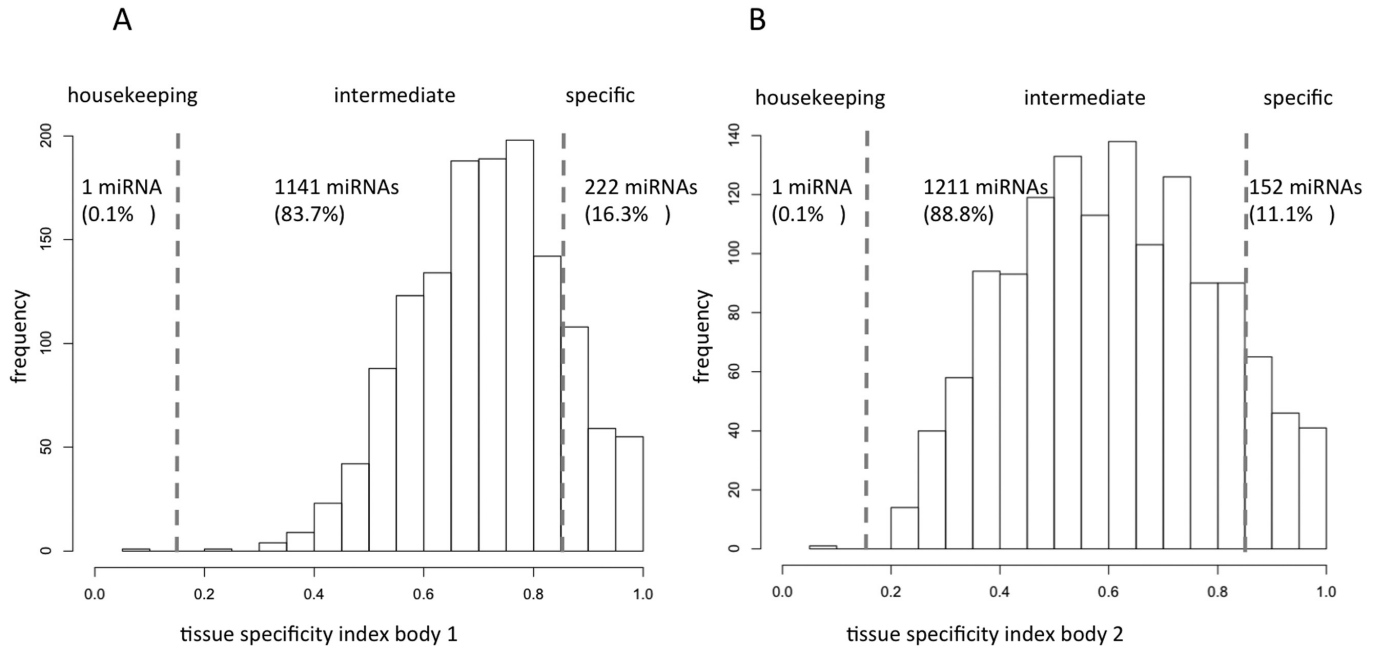

Figure 2. Histogram plot for the frequency of TSI of miRNAs in different tissues. Panel **A** represents TSI of the first, panel **B** of the second body. The vertical dotted lines correspond to the threshold originally proposed for defining housekeeping and specifically expressed miRNAs of <0.15 and >0.85. The same representation for VSN-normalized data is presented in Supplementary Figure S2.

NAs that are annotated with only one mature form, we only included those miRNAs that have two mature forms annotated and carried out the analyses in a paired manner (41% of the 1364 mature miRNAs were included). First, we investigated whether -3p or -5p mature forms are overall higher expressed. For both quantile- and VSN-normalized data, we calculated significantly higher expression of the -5p mature forms. The effects in VSN exceeded the quantile-normalized effects. Mature -5p forms were on average 21% higher expressed as compared to -3p forms (paired t-test *P*-value of 3.6×10^{-10}). To estimate whether the two mature forms are more or less specific for tissues, we calculated and compared the TSI values for the -3p and -5p forms. For both, TSI values based on VSN- and quantile-normalized data, we did not found significant differences between -3p and -5p forms ($P > 0.5$ in both cases). Having a detailed look at single miRNAs, we discovered that in all cases where -3p and -5p mature forms were tissue specific independent on the normalization technique the tissue patterns matched. The best matching profiles were found for hsa-miR-140, hsa-miR-378a, hsa-miR-509, hsa-miR-122, hsa-miR-124, hsa-miR-192 and hsa-miR-455. Only for one miRNA, miR-150, no significant correlation for -5p and -3p mature form was calculated (Supplementary Figure S4). The -3p form was specific for pancreas and the -5p form for stomach. All TSI values for -3p and -5p mature forms of quantile- and VSN-normalized data are available in Supplementary Table S4.

Inter-individual variations

In the previous analyses, we suggested that miRNAs are tissue specific. From two bodies it is impossible to extrapolate inter-individual variations within specific organs. In a first approach we searched for miRNAs that are overall higher or lower in all tissues of one of the two bodies, independent of the normalization technique. Two miRNAs, hsa-miR-548n and hsa-miR-548ap-5p, fulfilled these stringent criteria. Although these (and similarly differentially abundant miRNAs between both individuals) miRNAs had low TSI values and are not considered tissue specific the differences emphasize the importance of incorporating inter-individual variations.

We exemplarily analysed 16 lung tissue biopsies of 16 different individuals. Here, we expect miRNAs to be more homogenously expressed, leading to overall lower TSI values. For the quantile- and VSN-normalized data, we calculated significantly decreased TSI values in the individuals ($P < 10^{-16}$). The respective TSI values for biological replicates of lung tissue and the two bodies are presented in Supplementary Figure S5A (quantile normalized) and 5B (VSN normalized). These figures also indicate that few miRNAs have higher TSI in lung as compared to the overall TSI, i.e. variations between organs are smaller than variations between individuals. Inspecting the respective miRNAs, we found that they usually were specific for other organs than the lung and expressed to a very moderate limit in the lung. Here, already small variations lead to artificially high TSI values.

As the second example we downloaded expression values from 32 prostate tissues from the GEO (not affected tissues as part of a case-control cancer study, GSE76260). The TSI values were calculated for quantile- and VSN-normalized intensity values. Only the 625 miRNAs that were included in both studies were considered. In this analysis the variations between individuals were even lower as compared to

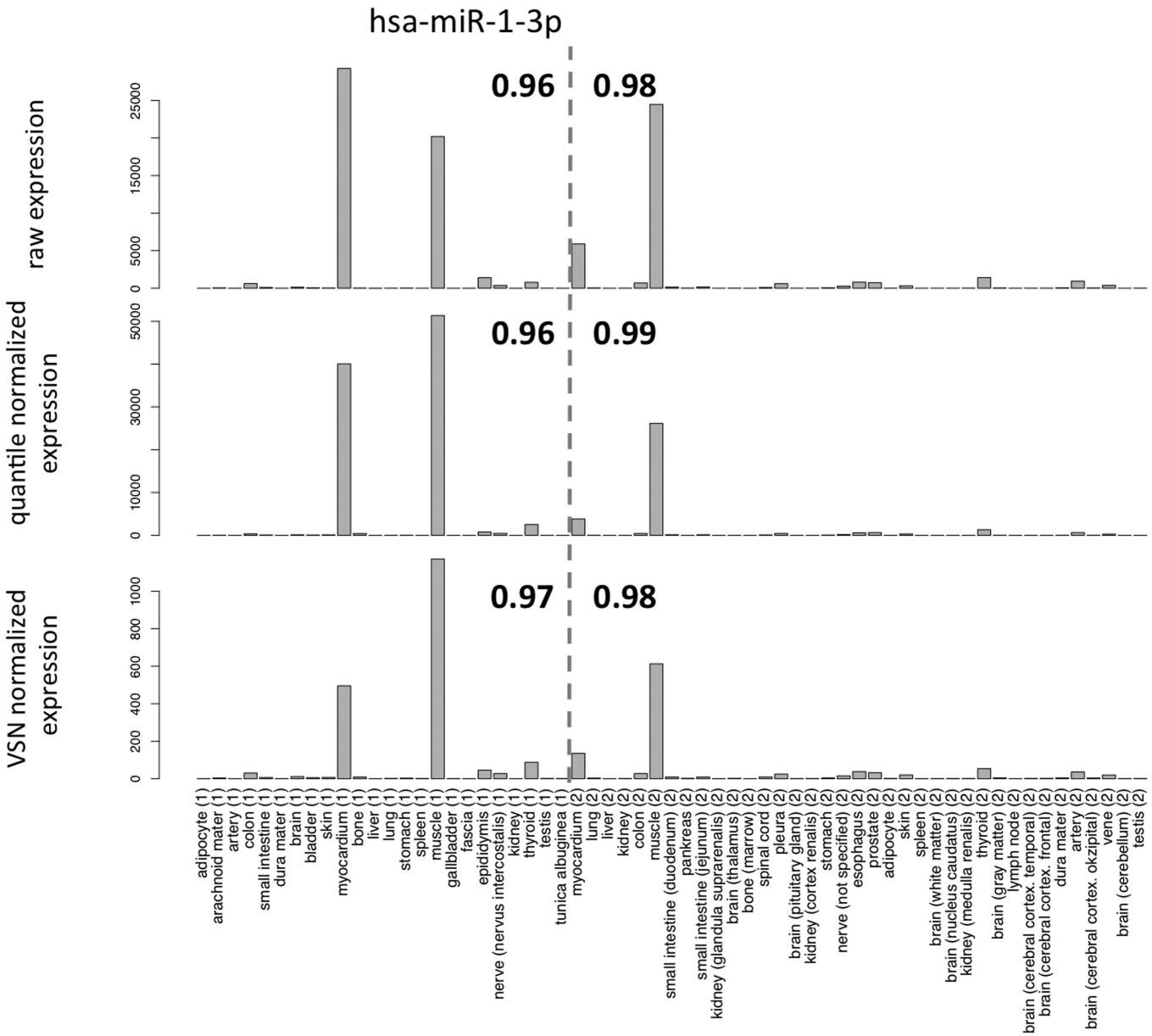

Figure 3. Bar plots for all 61 samples for miR-1–3p, the miRNA with highest overall TSI in the first and second body. The vertical dashed line separates the first from the second body. TSI values for both bodies are highlighted in the figure. The miRNA is high expressed in muscle and myocardium. Raw-, quantile- and VSN-normalized expression intensities for this miRNA match well across all different tissues.

the variations between organs. Again, TSI values were significantly lower for prostate tissue ($P < 10^{-16}$). The scatter plots are analogously to the lung tissues presented in Supplementary Figure S6. Also for the other tissues extracted from the GEO, which are also available on the tissue atlas web resource, lower TSI values were observed. In sum our results thus indicate that the inter-individual variations are smaller as compared to inter-organ variability.

Homology of tissue specificity in humans and rats

To addressed the question to what extend a tissue-specific abundance of the miRNA pattern is conserved between human and rodents, we matched the data of our study to data published in a recent study, which used the same miRNA platform (Agilent) (8). From all miRNAs expressed in our tissue collection, 230 matched in sequence identically between human and rat. Of the tissues included in the human and rat studies, 42 organs could be matched. For all these miRNAs and organs, we calculated the TSI values in human and rat, showing an overall correlation of 0.362 (P-value of 9×10^{-8}). To determine the significance of this finding, we additionally performed 1 million permutation tests, which showed an average correlation value of 0. While these results indicate an overall matching of miRNA abundances in humans and rats, the TSI does not acknowledge the origin of the miRNAs, i.e. a value of 1 for a rat miRNA may indicate

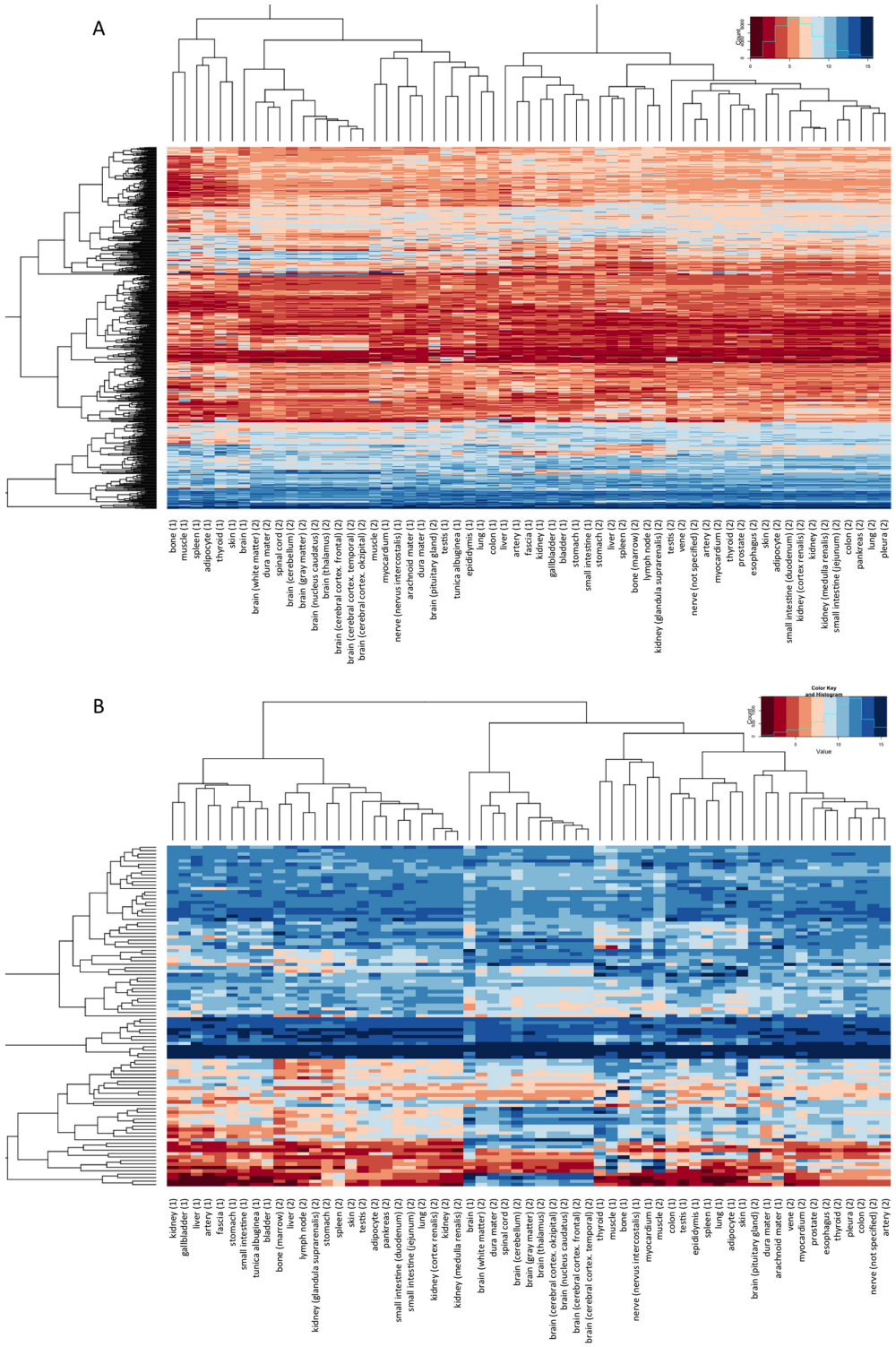

Figure 4. Hierarchical clustering of all tissues in both bodies. Log_2 transformed quantile normalized intensity values were used for clustering. The intensity value distribution is shown in the upper right corner of the figures. Panel **A** shows significantly expressed miRNAs, while panel **B** focuses on the 100 miRNAs with overall highest data variance. The respective representation for VSN-normalized data is presented in Supplementary Figure S3.

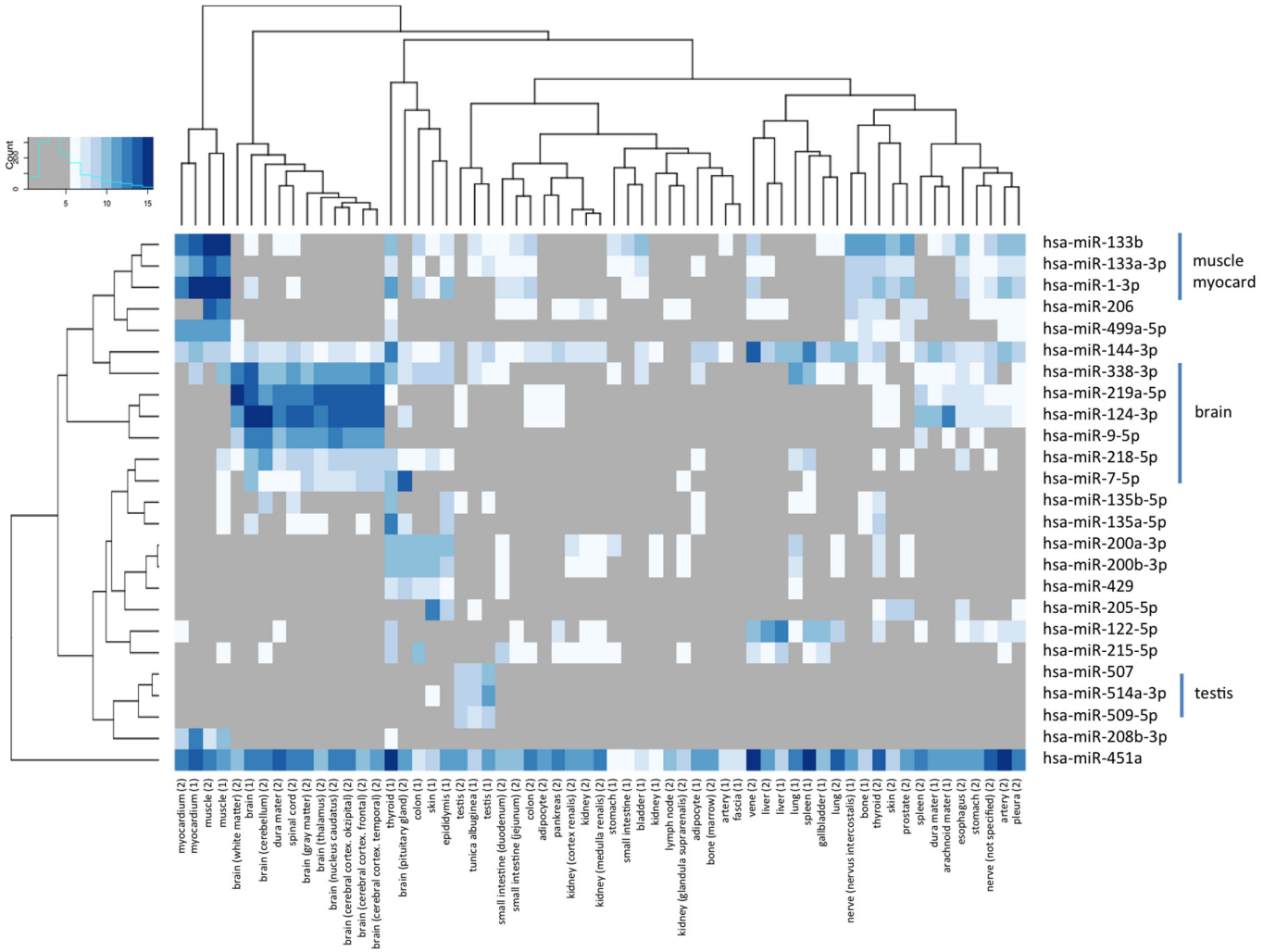

Figure 5. Heat map for the 25 miRNAs that have TSI values of >0.85 in both bodies. Log_2 transformed expression intensities of quantile normalized expression values are presented. To facilitate the interpretation of specific miRNAs in organs or organ groups low expressed miRNAs were greyed out (see also color distribution scheme in the upper left corner). The analysis highlights tissue-specific miRNAs that are exemplarily presented on the right-hand side of the plot, such as hsa-miR-1–3p that has already been described in Figure 3 as most specific miRNA overall.

specificity for spleen and for the same miRNA specificity for brain in humans. However, the overall correlation of the expression values of rat and human miRNAs was 0.361 ($P < 10^{-16}$), indicative of a significant matching of human and rat expression profiles. Similar to the results for humans in Figure 5, we clustered the miRNAs with high TSI values in human and rat. Altogether, we focused on very specific miRNAs: 54 miRNAs with TSI values exceeding 0.9 were considered. The resulting heat map where maximal rat and human miRNA expression was set to 100% to make both data sets comparable to each other is presented in Figure 7. In this analysis we did not observe a predominant clustering in humans and rats but a strong tendency of organs to cluster together. Examples of directly matching pairs include the spleen, myocardium, muscle, pancreas, kidney, liver, stomach, skin, brain or spinal cord. The miRNAs in this heat map matched the specific miRNAs in Figure 5 very well such as miR-133a-3p, and miR-133b for muscle and myocardium or miR-9–5p, miR-219a-5p, miR-7–5p and miR-124–3p for brain and spinal cord. Bar plots comparing each miRNA directly for specificity in tissues of rat and human are provided in the supplementary material.

DISCUSSION

As miRNAs emerge as important regulators of protein expression during tissue development and homeostasis, there is an increasing need for a standardized atlas of miRNA expression in multiple human tissues. Although there is ample evidence for differential miRNA expression in different human tissues, the majority of studies investigate differential expression in only one organ/tissue. Due to the different identification methods and normalization strategies, the results of these studies are not easily comparable limiting their value for comparison of miRNA expression in different tissues. The optimal human miRNA tissue atlas would be based on different fresh tissues each obtained from the same donor; different donors should be of different age and gender both of which are known to influence the miRNA

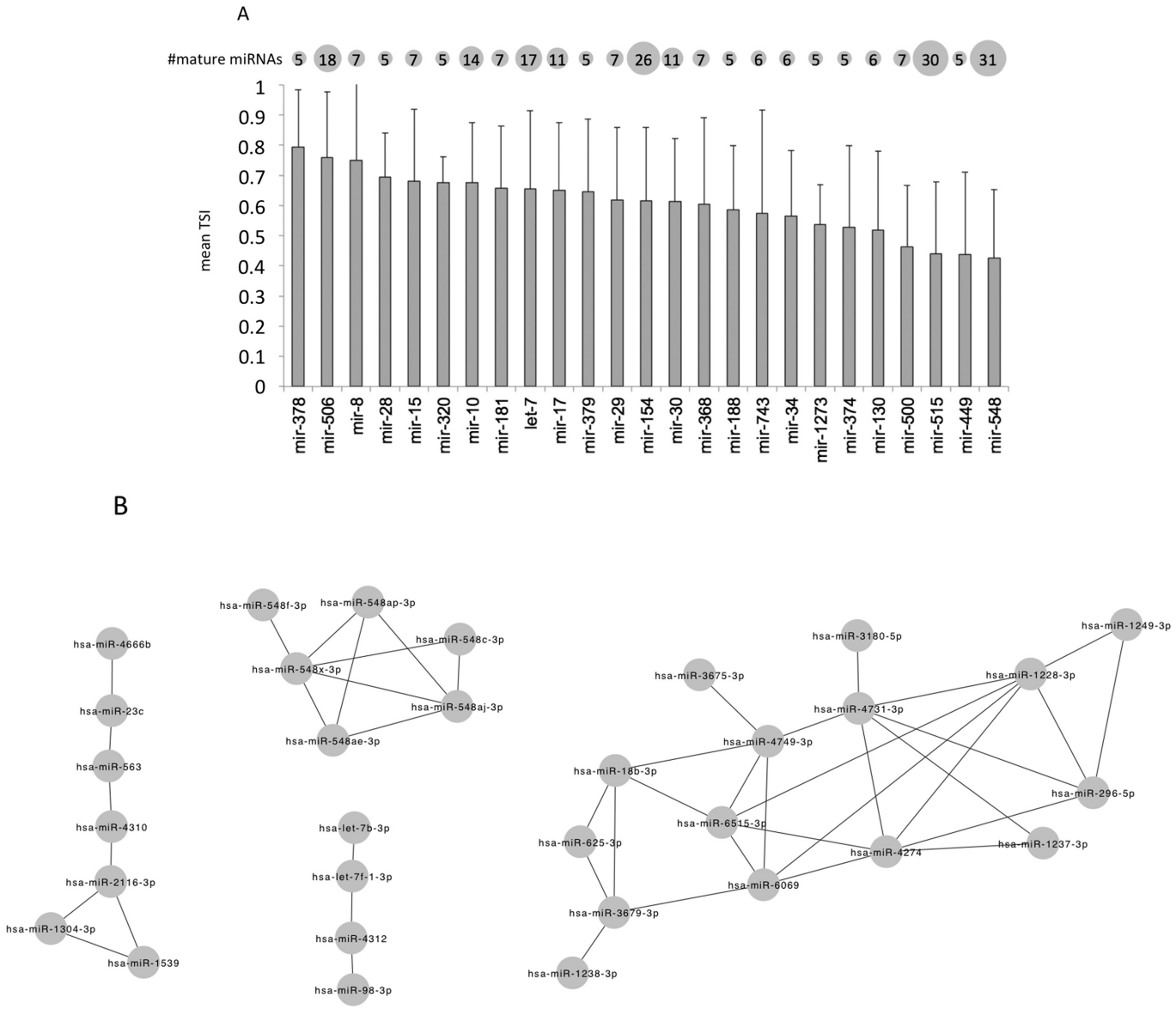

Figure 6. **A**: Average and standard deviation of TSI value in different miRNA families. For each miRNA family with at least five members the mean and standard deviation of all family members TSI is presented as bar plot. Families are sorted with decreasing average tissue specificity from left to right. Highest tissue specificity was observed for the miR-378 family, predominantly being specific for myocardium and muscle. The number of mature family members is shown above the columns with balloons, representing the family size. **B**: Co-expression network of miRNAs. Each miRNA pair connected by an edge has co-expression across all samples with Spearman correlation coefficient above 0.95.

pattern (12). As this ideal scenario is not possible in human studies, fresh biopsy material could be used for miRNA isolation with the advantage of yielding high-quality RNA. There are, however, several disadvantages: (i) biopsies will be mostly taken from patients with affected organs, (ii) high inter-individual differences can mask tissue-specific differences of miRNA abundances, (iii) a bias is likely introduced by multiple centres that are involved in tissue collections and (iv) samples of vital organs, e.g. thalamus, spinal cord or cerebellum, are not available. Alternatively, miRNAs can be isolated from tissues collected from the same individuals upon autopsy. The advantage of the latter approach is the availability of multiple tissues from the same individuals, even from vital organs, with the disadvantage of RNA degradation in the samples due to the storage duration of the body and the advanced age or the disease status of the body donors. In context of our tissue atlas, the main question is whether the differences in the abundance of miRNAs induced by post-mortem RNA degradation, which is different from *in-vitro* RNA degradation by UV light or heat, are higher than the differences between the tissues profiled. There is scant evidence for extended post-mortem stability of individual miRNAs (13,14). In case of whole miRNA tissue profiles, Ibberson et al. found that RNA degradation due to prolonged inadequate tissue storage has a random effect on miRNAs and compromises the reliability of miRNA

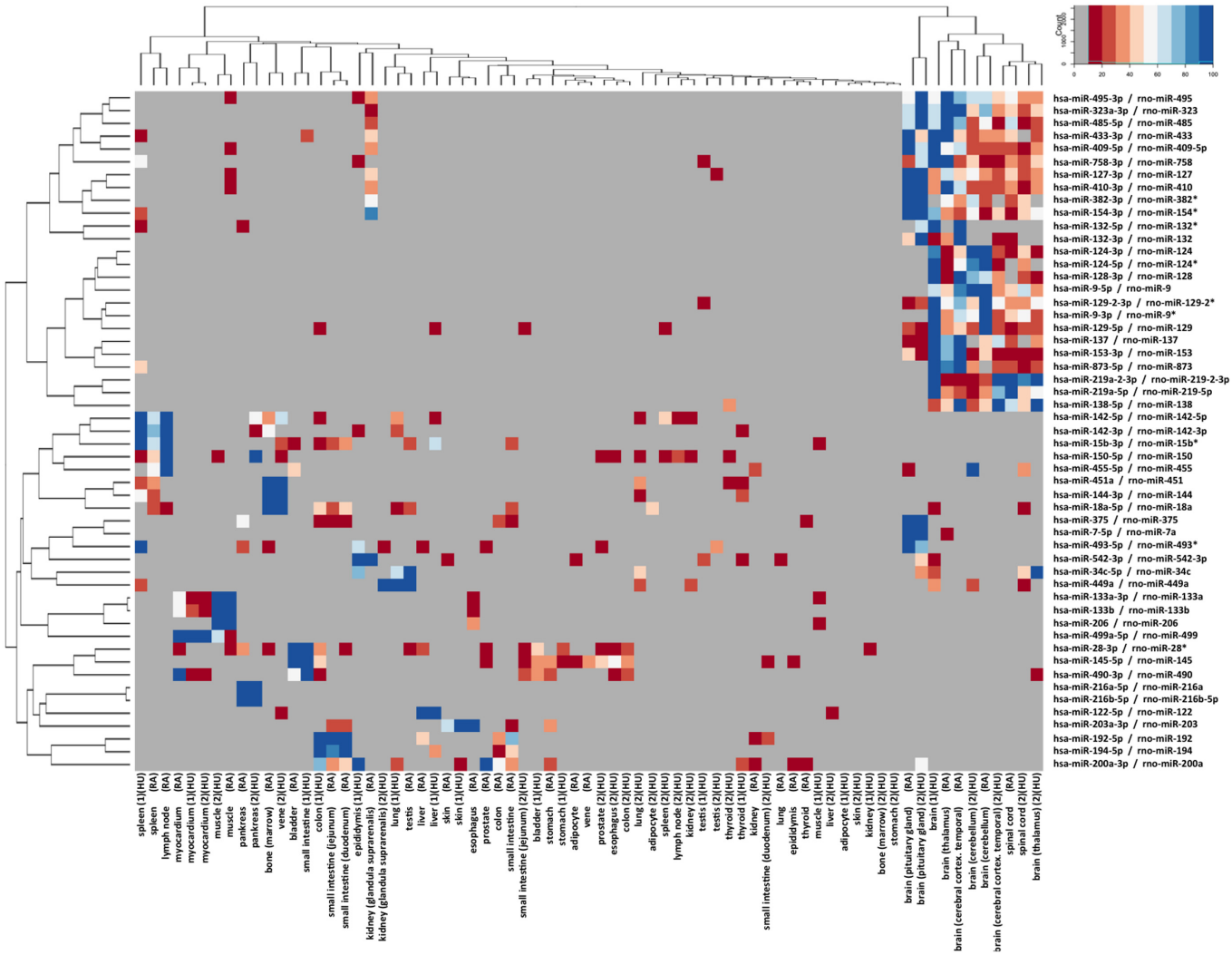

Figure 7. Conservation of tissue-specific expression of miRNAs in human and rat. Matching miRNAs (100% matching of mature miRNA sequence) from organ expression in rats and humans were calculated. For each miRNA in rats and humans the TSI was calculated and highly specific miRNAs were clustered. Since overall expression in humans and rats varied, the maximal intensity of each miRNA in the two organs was set to 100% and all other miRNAs were linearly scaled. All miRNAs with below 10% expression of maximal intensity are shown in grey to facilitate data interpretation (see also colour gradient presented in the upper right corner). On the right-hand side the human/rat miRNA identifiers are shown, below the heat map the matched tissues are presented (HU for human; RA for rat). For rat tissues the average intensity of replicated measurements is presented.

profiles, generating false positive deregulated miRNAs (15). But they also clearly state that 'even samples with the most degraded RNAs still preserve a tissue-specific miRNA signature'. This finding is in line with our observations in the present study. For lung and heart tissue we investigated short- and long-term degradation, highlighting an overall limited impact on the tissue specificity of miRNA profiles. Only very few miRNAs were affected at all. Given the data from two organs, we however cannot exclude the possibility that some tissue-specific miRNAs might be affected by degradation of the sample. We are also aware that the autopsy samples of the two male individuals provide only a snapshot of the full variability of miRNA expression. While we aim at adding more full body profiles we supported the data in the present study by tissue collections extracted from the literature (e.g. gastric and prostate tissues) and by own measurements (lung tissue).

We used a microarray platform for miRNA expression detection since this platform shows a high reproducibility as evidenced by the miRQC study (5). In our study, analysis of technical replicates of nine samples processed in different batches reached high correlation values above 0.986 for all samples. In previous studies, we observed a substantial bias introduced in Next Generation Sequencing (NGS) data by sample preparation of blood samples (10). However, NGS analysis would enable to detect presently unknown miRNAs as well miRNAs iso-forms that have demonstrated to target biological pathways in a cooperative manner (16). A key challenge with microarray data is normalization. Many techniques that are frequently applied such as variance stabilizing normalization or quantile normalization can have a substantial influence on the results. Quantile normalization e.g. assumes an overall similar distribution of all miRNAs. We thus performed the relevant analyses on raw data,

quantile- and VSN normalization. Irrespective of the normalization technique we found higher TSI values for miRNAs as, e.g. known from mRNAs (7). This result suggests that miRNA expression is more tissue specific as compared to mRNA expression.

The, as of now, most comprehensive study on tissue-specific miRNAs in humans was published by Landgraf et al. in 2007 (3). They sequenced 256 small RNA libraries from 26 different organ systems and cell types of humans and rodents, with ∼1000 clone each. The human samples included normal samples from 16 tissues most of them brain and reproductive tissues. They identified 340 mature human miRNAs including 33 novel miRNAs not listed in the miRBase version 9.1, which was the current version at the time of the study (17). For canonical miRNAs they found a high concordance of tissue-specific expression in humans and rodents. When we compared our data to a data set on 55 different rat tissues available at GEO database (8), we could confirm conserved tissue-specific expression of several miRNAs, including miR-133b, miR-124 and miR-9. Amongst others, Landgraf et al. detected tissue-specific expression of miR-122 in liver, of miR-9, miR-124 and miR 128a/b in brain, of miR-7, miR-375, miR-141 and miR-200a in pituitary gland and of miR-142, miR-144, miR-150, miR-155 and miR-223 in hematopoietic cells. Overall, our results correlated well with this data, confirming specific expression of miR-122, miR-9, miR-124 and miR-7 in the respective organs. Consistent with Landgraf's results, we found miR-122–5p as highest expressed miRNA in the liver of both bodies. Our study, however, also identified low expression of miR-122–5p in spleen, gall bladder and veins. MiR-124 (miR-124–3p) was identified as the third most specific miRNA in the nervous system by Landgraf et al. We observed expression of this miRNA in different areas of the brain but not in other tissues. For miR-144, we found highest expression in vein and spleen, consistent with the assumption of residual hematopoietic cells in these samples; additionally, we found high expression of this miRNA in thyroid. Of note, miR-144 has been found highly expressed in normal thyroid and downregulated in papillary thyroid carcinoma (18). We also found high expression of miR-1–3p, miR-133a-3p, miR-133b and miR-206 in myocard and muscle. These miRNAs are known as myomiRs that regulate key genes in muscle development (19,20). Additionally, we detected a highly specific expression of miR-205–5p, miR-514a-3p and miR-192–5p in skin, testis and colon samples of one of the bodies, respectively. MiR-205–5p that is highly expressed in melanocytes and downregulated in melanoma is inverse correlated with melanoma progression (21). MiR-514a-3p belongs to the miR-506 family; the mouse orthologue of miR-506, mmu-201, has been shown to be specifically expressed in reproductive tissues (3). A significant decrease in expression of miR-192–5p in colorectal cancer compared to normal mucosa has been reported (22).

The knowledge of the expression pattern of miRNAs in different tissues is essential for understanding normal development and disease development of the respective tissue. In addition, knowing the tissues that express specific miRNAs helps to develop a miRNA found in whole blood or serum into a biomarker for a specific disease. Elevated serum levels of liver-specific miR-122 have been detected in patients with drug induced liver injury, steatosis, hepatitis-B and -C infections and in patients with hepatocellular carcinoma (23–26). Elevated levels of circulating myomiRs, i.e. miR-1, miR-206 and miR-133a/b, have been proposed as biomarker for heart failure and different forms of muscle dystrophy, but are also elevated after half-marathon run (27–29).

In summary, we provide an atlas of miRNA expression in multiple human tissues. This atlas can be used as starting point for elucidation of the role of miRNAs in tissue development and tissue-specific diseases.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

We acknowledge the support of Siemens Healthcare.

FUNDING

Saarland University and Siemens Healthcare. Funding for open access charge: Saarland University and Siemens Healthcare; funded in part by FP7 project BestAgeing.
Conflict of interest statement. None declared.

REFERENCES

1. Petryszak,R., Burdett,T., Fiorelli,B., Fonseca,N.A., Gonzalez-Porta,M., Hastings,E., Huber,W., Jupp,S., Keays,M., Kryvych,N. *et al.* (2014) Expression atlas update–a database of gene and transcript expression from microarray- and sequencing-based functional genomics experiments. *Nucleic Acids Res.*, **42**, D926–D932.
2. Ponten,F., Jirstrom,K. and Uhlen,M. (2008) The human protein atlas–a tool for pathology. *J. Pathol.*, **216**, 387–393.
3. Landgraf,P., Rusu,M., Sheridan,R., Sewer,A., Iovino,N., Aravin,A., Pfeffer,S., Rice,A., Kamphorst,A.O., Landthaler,M. *et al.* (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell*, **129**, 1401–1414.
4. Leidinger,P., Backes,C., Meder,B., Meese,E. and Keller,A. (2014) The human miRNA repertoire of different blood compounds. *BMC Genomics*, **15**, 474.
5. Mestdagh,P., Hartmann,N., Baeriswyl,L., Andreasen,D., Bernard,N., Chen,C., Cheo,D., D'Andrade,P., DeMayo,M., Dennis,L. *et al.* (2014) Evaluation of quantitative miRNA expression platforms in the microRNA quality control (miRQC) study. *Nat. Methods*, **11**, 809–815.
6. Huber,W., von Heydebreck,A., Sultmann,H., Poustka,A. and Vingron,M. (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*, **18**(Suppl. 1), S96–S104.
7. Yanai,I., Benjamin,H., Shmoish,M., Chalifa-Caspi,V., Shklar,M., Ophir,R., Bar-Even,A., Horn-Saban,S., Safran,M., Domany,E. *et al.* (2005) Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics*, **21**, 650–659.
8. Minami,K., Uehara,T., Morikawa,Y., Omura,K., Kanki,M., Horinouchi,A., Ono,A., Yamada,H., Ohno,Y. and Urushidani,T. (2014) miRNA expression atlas in male rat. *Sci. Data*, **1**, 140005.
9. Edgar,R., Domrachev,M. and Lash,A.E. (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.*, **30**, 207–210.
10. Backes,C., Leidinger,P., Altmann,G., Wuerstle,M., Meder,B., Galata,V., Mueller,S.C., Sickert,D., Stahler,C., Meese,E. *et al.* (2015) Influence of next-generation sequencing and storage conditions on miRNA patterns generated from PAXgene blood. *Anal. Chem.*, **87**, 8910–8916.

11. Backes,C., Sedaghat-Hamedani,F., Frese,K., Hart,M., Ludwig,N., Meder,B., Meese,E. and Keller,A. (2016) Bias in high-throughput analysis of miRNAs and implications for biomarker studies. *Anal. Chem.*, **88**, 2088–2095.
12. Meder,B., Backes,C., Haas,J., Leidinger,P., Stahler,C., Grossmann,T., Vogel,B., Frese,K., Giannitsis,E., Katus,H.A. *et al.* (2014) Influence of the confounding factors age and sex on microRNA profiles from peripheral blood. *Clin. Chem.*, **60**, 1200–1208.
13. Nagy,C., Maheu,M., Lopez,J.P., Vaillancourt,K., Cruceanu,C., Gross,J.A., Arnovitz,M., Mechawar,N. and Turecki,G. (2015) Effects of postmortem interval on biomolecule integrity in the brain. *J. Neuropathol. Exp. Neurol.*, **74**, 459–469.
14. Lv,Y.H., Ma,K.J., Zhang,H., He,M., Zhang,P., Shen,Y.W., Jiang,N., Ma,D. and Chen,L. (2014) A time course study demonstrating mRNA, microRNA, 18S rRNA, and U6 snRNA changes to estimate PMI in deceased rat's spleen. *J. Forensic Sci.*, **59**, 1286–1294.
15. Ibberson,D., Benes,V., Muckenthaler,M.U. and Castoldi,M. (2009) RNA degradation compromises the reliability of microRNA expression profiling. *BMC Biotechnol.*, **9**, 102.
16. Cloonan,N., Wani,S., Xu,Q., Gu,J., Lea,K., Heater,S., Barbacioru,C., Steptoe,A.L., Martin,H.C., Nourbakhsh,E. *et al.* (2011) MicroRNAs and their isomiRs function cooperatively to target common biological pathways. *Genome Biol.*, **12**, R126.
17. Griffiths-Jones,S., Grocock,R.J., van Dongen,S., Bateman,A. and Enright,A.J. (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.*, **34**, D140–D144.
18. Swierniak,M., Wojcicka,A., Czetwertynska,M., Stachlewska,E., Maciag,M., Wiechno,W., Gornicka,B., Bogdanska,M., Koperski,L., de la Chapelle,A. *et al.* (2013) In-depth characterization of the microRNA transcriptome in normal thyroid and papillary thyroid carcinoma. *J. Clin. Endocrinol. Metab.*, **98**, E1401–E1409.
19. Callis,T.E., Chen,J.F. and Wang,D.Z. (2007) MicroRNAs in skeletal and cardiac muscle development. *DNA Cell Biol.*, **26**, 219–225.
20. Thum,T., Catalucci,D. and Bauersachs,J. (2008) MicroRNAs: novel regulators in cardiac development and disease. *Cardiovasc. Res.*, **79**, 562–570.
21. Liu,S., Tetzlaff,M.T., Liu,A., Liegl-Atzwanger,B., Guo,J. and Xu,X. (2012) Loss of microRNA-205 expression is associated with melanoma progression. *Lab. Invest.*, **92**, 1084–1096.
22. Karaayvaz,M., Pal,T., Song,B., Zhang,C., Georgakopoulos,P., Mehmood,S., Burke,S., Shroyer,K. and Ju,J. (2011) Prognostic significance of miR-215 in colon cancer. *Clin. Colorectal Cancer*, **10**, 340–347.
23. Akamatsu,S., Hayes,C.N., Tsuge,M., Miki,D., Akiyama,R., Abe,H., Ochi,H., Hiraga,N., Imamura,M., Takahashi,S. *et al.* (2015) Differences in serum microRNA profiles in hepatitis B and C virus infection. *J. Infect.*, **70**, 273–287.
24. Krauskopf,J., Caiment,F., Claessen,S.M., Johnson,K.J., Warner,R.L., Schomaker,S.J., Burt,D.A., Aubrecht,J. and Kleinjans,J.C. (2015) Application of high-throughput sequencing to circulating microRNAs reveals novel biomarkers for drug-induced liver injury. *Toxicol. Sci.*, **143**, 268–276.
25. Pirola,C.J., Fernandez Gianotti,T., Castano,G.O., Mallardi,P., San Martino,J., Mora Gonzalez Lopez Ledesma,M., Flichman,D., Mirshahi,F., Sanyal,A.J. and Sookoian,S. (2015) Circulating microRNA signature in non-alcoholic fatty liver disease: from serum non-coding RNAs to liver histology and disease pathogenesis. *Gut*, **64**, 800–812.
26. Xu,J., Wu,C., Che,X., Wang,L., Yu,D., Zhang,T., Huang,L., Li,H., Tan,W., Wang,C. *et al.* (2011) Circulating microRNAs, miR-21, miR-122, and miR-223, in patients with hepatocellular carcinoma or chronic hepatitis. *Mol. Carcinog.*, **50**, 136–142.
27. Akat,K.M., Moore-McGriff,D., Morozov,P., Brown,M., Gogakos,T., Correa Da Rosa,J., Mihailovic,A., Sauer,M., Ji,R., Ramarathnam,A. *et al.* (2014) Comparative RNA-sequencing analysis of myocardial and circulating small RNAs in human heart failure and their utility as biomarkers. *Proc. Natl Acad. Sci. USA*, **111**, 11151–11156.
28. Gomes,C.P., Oliveira-Jr,G.P., Madrid,B., Almeida,J.A., Franco,O.L. and Pereira,R.W. (2014) Circulating miR-1, miR-133a, and miR-206 levels are increased after a half-marathon run. *Biomarkers*, **19**, 585–589.
29. Cacchiarelli,D., Legnini,I., Martone,J., Cazzella,V., D'Amico,A., Bertini,E. and Bozzoni,I. (2011) miRNAs as serum biomarkers for Duchenne muscular dystrophy. *EMBO Mol. Med.*, **3**, 258–265.