Chapter 8
The Analysis of MicroRNAs in Stem Cells

Loyal A. Goff[1], Uma Lakshmipathy[2], and Ronald P. Hart[1*]

Abstract MicroRNAs represent a newly-discovered class of regulatory molecules that has been demonstrated to be required for stem cell function. Methods for measuring unique microRNAs are particularly useful in classifying stem cells or for studying mechanisms underlying their differentiation. Furthermore, straightforward bioinformatic and statistical methods are useful in investigating large sets of data to formulate hypotheses or identify microRNAs associated with a specific effect or phenotype. We present an overview of microRNA biology, detection techniques including microarrays, as well as methods for analyzing the resulting data in the context of stem cell function.

Keywords: microRNA, small RNA, Stem cell, microarray, deep sequencing, sample preparation, data analysis, biclustering.

8.1 Background

No current discussion of stem cell maintenance, regulation or differentiation can be complete without mention of the largest class of tiny regulators of gene expression, microRNAs. The relatively recent discovery of this class of small non-coding RNAs has turned many common assumptions regarding cellular networks on end. Various types of small, non-coding RNAs exist as modulators of gene expression, affecting transcription rate [50, 85], heterochromatin formation [107, 130], transposon silencing [133], mRNA stability [13, 46, 160], and mRNA translation into functional proteins [101, 102, 120]. MicroRNAs represent a class of endogenous

[1] W.M. Keck Center for Collaborative Neuroscience, and the Rutgers Stem Cell Research Center, Rutgers University, Piscataway, NJ, USA

[2] Regenerative Medicine, Invitrogen, Inc., Carlsbad, CA, USA

*Corresponding Author:
604 Allison Rd Rm D251, Piscataway, NJ 08854, USA
e-mail: rhart@rci.rutgers.edu

Y. Shi, D.O. Clegg (eds.) *Stem Cell Research and Therapeutics*,

© Springer Science+Business Media B.V. 2008

Shi_Ch08.indd 141 5/30/2008 5:59:55 PM

genes whose primary role appears to be post-translational regulation of specific target mRNAs [3, 11]. MicroRNA genes are encoded both as intergenic transcripts as well within intronic sequences. The concept of intronic microRNAs alone demonstrates that there is a greater compression of genetic information contained within our genomes than previously understood, and the identification of over 1,000 human microRNA sequences to date suggests that there is indeed more regulatory complexity than previously appreciated.

What exactly are microRNAs? While a mature microRNA can arise from an independent transcript, a poly-cistronic cluster, or an intronic sequence, each of these produces a stem-loop RNA precursor sequence consisting of ~100 paired bases. These stem loops then become the substrate for the multi-protein "Microprocessor" complex [37, 51]. Association with this complex confers binding by the dsRNA binding enzyme Pasha and enzymatic cleavage of the hairpin from the primary transcript by the nuclear RNase III enzyme Drosha [37, 55, 56, 91, 168, 169]. The cleaved hairpins are then exported from the nucleus via Exportin-5 and, once in the cytoplasm, associate with the miRNP (microribonucleoprotein) complex. The proteins within this complex include Dicer, a second RNase III enzyme that removes that loop from the hairpin sequence, and a helicase to separate the resulting duplex and remove the microRNA* strand (the strand complementary to the mature microRNA). The result is a 'primed' ribonuclear complex capable of targeting a specific sequence. This specificity is conferred by the presentation of the now single-stranded ~21mer mature microRNA, which will recognize the 3′ untranslated region (3′ UTR) of a target mRNA. At our present level of understanding, there are several mechanisms of action that might occur at this point. A bound miRNP complex demonstrating a high degree of complementarity is capable of activating its "Slicer" activity resulting in the cleavage of the mRNA transcript at the site of miRNP binding [100, 163]. This ultimately results in the degradation of the transcript. Additionally, mRNA degradation has been observed through 5′ de-capping mechanisms [13, 127] as well as rapid de-adenylation [46, 160]. More commonly with respect to mammalian microRNAs, the interaction of a primed miRNP and its specific target results not in the degradation of the mRNA but rather in the repression of translation. Again, several different mechanisms for translational repression have been observed. Bound microRNAs are capable of directing target mRNAs to specific sub-cellular locations known as P-bodies [13, 95, 101, 102, 120, 127], or simply stalling protein production through direct hindrance of ribosomes [80, 118] or interfering with translation initiation [68, 121]. Estimates suggest there are ~400 microRNA genes in each invertebrate species, and ~1,000–1,500 genes in mammals [93, 97], with some groups predicting as many as 10,000–20,000 microRNA genes per genome [109]. The widespread impact of this new layer of gene regulation is also becoming more apparent in that several groups estimate anywhere from ~30% to 95% of the genome may be targets for microRNAs [93, 109]. This type of gene control represents a novel regulatory mechanism, and is predicted to affect many crucial cellular processes and developmental programs, including neurogenesis.

Only a few validated target mRNAs have been identified in animals. This information, combined with correlated tissue expression data and functional analyses, highlights some of the important roles for microRNAs. microRNAs have been shown to play a role in numerous cancers [1, 23, 25, 26, 38, 42, 52, 59–61, 64, 70, 81, 84, 129, 131, 142, 149], cardiac hypertrophy/failure [149], and several other disorders. Additionally, cellular processes such as fat metabolism [40, 161], insulin regulation [122, 123], apoptosis [7, 29, 34, 162], cell cycle regulation [58, 95, 135], maternal-zygotic transition [46, 110, 156], viral defense [35, 91], axis specification/patterning [45, 57, 63, 71, 104], tissue formation [32, 43, 115, 164], as well as stem cell specification and differentiation [4, 16, 21, 32, 39, 58, 65, 66, 79, 92, 115, 126, 135, 137, 138, 140, 159, 171] have all been associated with microRNA activity.

Shortly after the identification of microRNAs, tissue surveys were conducted to assess the potential impact of these small inhibitors [6, 8, 83, 117, 145]. It became immediately apparent that most microRNAs are highly tissue restricted, with only a slight overlap between tissues for any given microRNA. A few microRNAs, such as the let-7 family, appear to be ubiquitously expressed in all tissue types [8, 117, 145], suggesting a role in regulating the more basal, and therefore more prevalent processes within the cell, such as cell-cycle regulation [70]. More commonly, a set of tissue specific microRNAs are associated with specific cellular functions. This correlation, while only predictive of a microRNA's role, is strengthened by associated functional analysis of microRNA activity via over-expression [70, 123, 148] and inhibition studies [39]. For example, antisense targeting of miR-122, a liver specific microRNA, led to dysregulation of lipid metabolism in the liver [40]. Using similar approaches of combining tissue-specific expression data with functional assays has led to a greater understanding of the impact of microRNAs in the cell, without the need for specifically identifying valid mRNA targets. Using this approach, groups have ascribed general functions for given microRNAs in context, such as the requirement for miR-125b for the proliferation of differentiated cells [92], miR-143 for the regulation of adipocyte differentiation [39], and numerous others [79, 146]. It would be shortsighted at this point to restrict the potential roles that microRNAs may be performing in the cell as new roles are being unraveled at an accelerated rate. The diversity in the already established roles for microRNAs demonstrates that this class of small regulatory RNA molecules plays an integral role in numerous biological pathways and suggests that they will play an important role in other cellular processes including differentiation.

One of the more interesting roles suggested for these small inhibitors of translation is the regulation and specification of stem cells. Several studies have attempted to determine the global role of microRNAs in development by selective knockdown of required components of the microRNA/RNAi pathway [17, 73]. A few groups have determined that Dicer, the RNase III enzyme responsible for processing microRNAs, and therefore required for microRNA activity, is required for murine cell differentiation and specification [17, 73, 114]. Evidence shows that a Dicer-1 null mutant mouse was embryonic lethal due to a depletion of stem cells [17], as well as demonstrating a failure of existing stem cells to adequately differentiate [73].

Shi_Ch08.indd 143 5/30/2008 5:59:55 PM

Interestingly, a similar study was conducted in zebrafish where it was shown that embryos with a maternal-zygotic Dicer mutant are capable of stem cell maintenance and differentiation but are defective in patterning, morphogenesis, and organogensis, suggesting that the role of microRNAs in stem cell regulation may have changed dramatically during the course of evolution [45]. A mouse knockout of the DGCR8 (Pasha) gene, required for the recognition and accurate processing of microRNA precursors, demonstrated the requirement of functional mature microRNAs for appropriate differentiation of ES cells. It was shown that ES cells lacking a functional DGCR8 gene failed to completely differentiate and retained the pluripotency markers Oct4 and Nanog, despite the onset of selected differentiation markers as well [151]. Furthermore, knockout mice lacking Argonaute2 (Ago2), the catalytic component of the RISC complex, exhibited severe defects in neural development, including the failure to close neural tube [100]. These experiments highlight the critical, if not yet well understood role that microRNAs play during stem cell development.

Early on, it was noted that stem cells express unique populations of microRNAs that were not present in any adult tissues [46, 65, 66, 110, 138], some of which additionally appear to be species specific. A conserved eutherian microRNA cluster is expressed exclusively in undifferentiated stem cells and is immediately downregulated upon induction of differentiation [65, 66]. Since microRNAs are hypothesized to have a predominately repressive role, it is reasonable to speculate that these microRNAs are responsible for maintaining a stem-like state through repression of pro-differentiation factors. A similar group of microRNAs, although with distinctly unique sequences and genomic locations, is evident in differentiating human embryonic stem cells [138]. The presence of these embryonic stem cell-specific microRNAs, and their clearance during differentiation, suggests a role in restricting cell differentiation. In contrast, new populations of tissue specific microRNAs are coordinately induced during differentiation and specification of stem cells [31, 32, 45, 74, 79, 83, 86, 89, 111, 140, 157]. Conserved microRNAs miR-1 and miR-206 are both induced during, and are required for, muscle cell differentiation and specification in mammals [32, 116] or birds [140]. Expression of miR-181 in hematopoietic stem cells is associated with an increase in B-cell specification [31], while other hematopoietic microRNAs (miR-142, and miR-155) are also induced during blood cell maturation [31, 125, 137]. Finally, subsets of microRNAs can be used to classify, for example, differences between embryonic stem cells, embryoid bodies, and embryonic carcinoma [89]. The requirement for microRNA activity during development and the influence that specific microRNAs have upon differentiation strengthen the argument that microRNAs are required for establishing and perhaps maintaining a differentiated state.

Without a doubt, microRNAs are present and active in both stem cell maintenance and differentiation. But before we are able to fully comprehend the roles of these small molecules, we must be familiar with the current methods for detection, analysis, and contextual interpretation of microRNAs. We present here an overview of current techniques in microRNA expression profiling, and suggest several possible workflows for analysis and interpretation of microRNA expression data.

8.2 Available Tools for Detecting MicroRNAs

Shortly after the realization that microRNAs were both abundant and ubiquitous regulators within the cell, many standard techniques were adapted to accommodate these new molecules. The first microRNAs identified were visualized by standard Northern blotting. A commonly-used and reliable technique, the Northern blot was easily deployed and required little adaptation to detect microRNAs. It unfortunately suffered from the short-comings of being fairly low throughput and time-consuming. Shortly after this, standard and quantitative real-time PCR (qPCR) were modified for identification and detection of microRNAs [30]. In general, qPCR techniques have been shown to yield the greatest dynamic range, improved specificity, and increased sensitivity in microRNA detection assays. In one case, qPCR allowed detection of microRNAs from single cultured neurons or laser captured somatodendritic compartments [87]. qPCR is a moderately high-throughput assay allowing rapid validation of a broader number of microRNAs than, for example, Northerns.

With the appearance of microRNA microarrays [6, 8, 48, 96, 111, 117, 145], however, a large number of microRNA genes could be assayed in parallel and true surveys or expression analyses could be conducted [6, 8, 83, 88, 117, 134, 145]. Several platforms emerged in rapid fashion [5, 8, 12, 18, 23, 24, 39, 47, 82, 96, 97, 99, 111, 139], and many commercial sources of microRNA arrays are readily available (Table 8.1); each with their own advantages and disadvantages. Most of these are based on the public list of microRNAs found in the miRBase database maintained at the Sanger Institute (http://microrna.sanger.ac.uk) [53]. A select few include probes for predicted microRNA genes or additional microRNAs that have not yet been indexed by the Sanger registry. Most of the commercial microRNA arrays are relatively quick to release updated probe sets as novel microRNAs are released in miRBase. Early considerations in probe design for microRNA microarrays focused mainly on the problem of variable melting temperatures (T_m) across microRNAs. To ensure an adequate and consistent signal during an array experiment, it is ideal for the probes to have a relatively narrow T_m range. With fairly short sequences from which to design probes, limited strategies are available to accomplish this goal. Solutions included logical sequence truncation or increasing the stringency of hybridization in one of several ways. With most of the array-based methods, it is difficult to claim resolution of specific microRNAs within 1 nt of the probe sequence since the melting temperatures are quite low compared with the longer probes often used for mRNA detection [48]. However, higher specificity can be achieved using direct labeling of microRNAs to obtain RNA:DNA hybridization (Ncode™) or by LNA oligo probes (mirCURY™), which enhance base stacking and phosphate-backbone reorganization, resulting in an increased thermal stability. Other detection techniques include ELISA-like or bead-anchored hybridizations using probes and labeling similar to array methods to perform high-throughput analyses on robotic liquid handling systems [114]. We focus this chapter on the use of any of the available microRNA microarray platforms for high-throughput analysis of microRNA expression.

Table 8.1 Commercial sources of microRNA microarrays

Product name	Source	References
Human miRNA Microarray	Agilent	[150]
mirVana™	Ambion/Applied Biosystems	[36, 170]
Species Specific MicroRNA Arrays	CombiMatrix	
MirCURY™	Exiqon, Inc.	[28, 112]
GenoExplorer™	GenoSensor Corp.	[9, 33, 154]
NCode™	Invitrogen, Inc.	[47, 72]
Human microRNA Microarray	LC Sciences	[67, 144]

Validation of microarray results originally depended on Northern blots but recently qPCR has become the method of choice. However, while qPCR may be a more convenient method for validation of a small to moderate number of micro RNAs, the microarray quality control project (MAQC) [136] determined that the best validation of an array experiment is to repeat the experiment on a separate array platform. Attention should be given in the project design phase to adequately prepare for some basic form of validation (i.e., ensure adequate material is available and that an acceptable method for validation has been considered).

8.3 Experimental Design Considerations

A microarray study is only as good as its experimental design. The amount of time spent planning and preparing an assay will pay off in the form of easily interpretable data, better quality results, and, ideally, straightforward answers to the experimental question. Consider what it is that you are interested in uncovering with an array study. A general survey alone to merely identify microRNAs present in a tissue sample will probably yield too little information and will most likely not produce results acceptable for publication. Conversely, a complex, broad, and unguided assay can be very confusing and cloud meaningful results in a sea of data.

A differential expression study can suggest whether or not microRNAs are being regulated (via any number of mechanisms) between two or more conditions. The number of conditions of interest will most likely determine the design of an experiment. A simple comparison of two conditions, for example, treated vs. control, can be conducted most effectively with a series of replicate two-color arrays. A more complex experimental design, such as a time course, multiple treatment conditions, or a multivariate study will require more careful design considerations.

Regardless of the scale of your experiment, replicate samples should be employed to capture and account for biological variability. Obviously, the more replicates that are used, the more statistical power is gained and the more confidence can be expressed in reporting results. The trade-off has traditionally been that more replicates require a significantly larger cost. This should not be so limiting since the relative cost per sample has decreased substantially in the past few years,

and will continue to decrease as new technology and new competition emerges. True biological replicates should be balanced with respect to arrays and dyes. We recommend an absolute minimum of three replicates per sample, but encourage the investigator to sample as many as is economically feasible for any given experiment.

8.4 MicroRNA Preparation and Handling

As with any RNA work, care and consideration must be given to ensure a clean, sterile, and nuclease-free environment. RNA is considerably more susceptible to degradation than DNA, and we find that degradation of the smaller microRNA generally is associated with degradation of higher molecular weight cellular RNAs such as ribosomal RNAs. Care should be taken to wear gloves when handling isolated RNA to protect your sample from RNAses found on the skin. RNA work should be conducted in a dedicated and clean space that is routinely treated with a nuclease inhibitor (e.g. RNaseZap, Ambion Cat #0611001A). If you begin to notice increased degradation of your isolated RNA, or generally reduced signal intensities in consecutive experiments, a thorough cleaning of your workspace is recommended.

An early realization was that the majority of labs involved in RNA work were commonly discarding RNA comprising less than ~100 bases in length. This was assumed to be primarily degradation products resulting from the RNA isolation techniques themselves. A widespread technique involved the ethanol-mediated binding of RNA to silica gel cartridges (for example, Qiagen RNeasy™). This allowed for the retention of large RNA molecules and the exclusion of anything passing through the cartridge, including weakly binding smaller RNA molecules. This method was easily modified by the manufacturer for retention of microRNA sequences by increasing the concentration of ethanol to drive a stronger affinity for the silica. While this did increase the yield of microRNAs, the increase in ethanol concentration often carried over to downstream applications potentially affecting reaction efficiencies. After much trial and error, we recommend a standard Trizol (Invitrogen Cat # 15596-018) RNA isolation followed by ammonium acetate/ethanol precipitation. If required, a carrier such as linear acrylamide (Ambion Cat #AM9520) can be used to increase yield and help precipitate the RNA. The result is an ultra-pure total RNA preparation that contains RNA of all sizes including microRNAs. Downstream applications may require that the microRNAs be isolated from the total RNA population. Several commercial products, including the miR-Vana™ Kit (Ambion, Austin, TX) or the PureLink™ miRNA Isolation Kit (Invitrogen, Carlsbad, CA), have been developed that will allow size fractionation based on selective column binding affinities or gel electrophoresis. We propose a simpler method of size fractionation via filtration. Centrifuge filter devices are a relatively inexpensive method for size selection and have demonstrated good isolation and separation properties. Begin by passing a sample of total RNA preparation through a ~100,000 molecular weight cutoff (MWCO) filter (Microcon YM-100 Cat# 42412). The microRNAs and any other small RNA (≤100 bp) will pass

through the filter and the larger RNA (mRNA, rRNA) will remain as retentate. The flow-through contains microRNA. The mRNA may be recovered by inverting and centrifuging the filter. Since the microRNA will co-purify with any low molecular weight contaminants, an additional clean-up procedure step is to concentrate the sample on a 3,000 MWCO filter (YM-3) to remove salts, phenol, or other impurities.

There are several rapid and simple assays that can be conducted on isolated microRNA that will test the quality and quantity of the total RNA or microRNA. A common technique to quantify an RNA preparation is to measure the absorbance at 260 nm (A_{260}) with a low-volume spectrophotometer, such as the Nanodrop ND-1000, which requires only 1–2 µl of sample. Typically, the absorbance is measured, corrected for dilution, and multiplied by a constant conversion factor to determine the concentration. The conversion factor for RNA is typically 40 µg/ml per A_{260} unit. This constant is derived from both the average molecular weight of RNA and the extinction coefficient. Since microRNAs have a significantly shorter sequence than the average RNA, this coefficient is insufficient. The approximation of the constant for microRNA that should be used to calculate concentration is 33 µg/ml per A_{260} unit. This is an important consideration only when microRNAs are isolated separately from total RNA.

In addition to microRNA quantity, quality of the preparation can be inferred by measuring the A_{230} and A_{280}. Since typical contaminating components absorb light at these two wavelengths (e.g., proteins at ~280 nm, ethanol at ~230 nm) the standard ratios A_{260}/A_{280} and A_{260}/A_{230} can be used as a measure of RNA purity. A "clean" preparation typically produces A_{260}/A_{280} and A_{260}/A_{230} ratios ≥ 2.0 (when measured at pH 8). If your sample produces values in either of these ratios < 1.8, we recommend additional cleanup steps to remove contaminants carried over from the RNA isolation.

While the spectrophotometer is a useful device to quantify RNA samples and determine their purity, it cannot determine the quality of the RNA itself (i.e., how intact is the RNA). High molecular weight RNA or total RNA fractions are typically assessed by gel electrophoresis. This is not realistic for microRNA samples, which regularly do not have enough mass to spare for a gel analysis. A practical solution is capillary electrophoresis. We recommend the Agilent Bioanalyzer 2100 system (Agilent, Santa Clara, CA) for several reasons. There are two available RNA quality assays that each measure different concentrations of input RNA material, although the quantity measurements are listed by the manufacturer as accurate ±50%. Additionally, the small RNA bioanalyzer kit enables the characterization of microRNA down to the picogram/microliter level [105]. The calculated RNA Integrity Number (RIN) provides a consistent and standard metric for the evaluation of total RNA quality. In the absence of this system, we recommend traditional agarose gel electrophoresis of the high molecular weight RNA fraction, or unfractionated total RNA, as a proxy for the assessment of microRNA quality within the same sample.

8.5 Probe-Level Interpretation of MicroRNA Array Data

Most commercial microRNA microarray platforms focus exclusively on the microRNAs derived from the most common model organisms; human, mouse, and rat. At the time of writing this chapter, the NCode™ platform from Invitrogen was designed to include probes for the above named species and additionally, probes for zebrafish, *Drosophila*, and the nematode *C. elegans* [48]. Since multiple microRNAs appear to share strong evolutionary origins and are highly conserved across multiple lineages, we can use the information content available to expand both the applications of these arrays, as well as our understanding of the expression of microRNAs across multiple species.

The high level of interspecies conservation of a large set of microRNAs means that the probes designed against human microRNAs, for example, may be useful for identifying novel microRNAs from similar vertebrate species. Furthermore, since most of the microRNA array platforms cannot distinguish between two molecules with 1 bp difference between them, those microRNAs that have diverged ≤1 bp should be detected by that probe. Through the course of our array design, we have been able to successfully map all known microRNAs from human, mouse, rat, zebrafish, *Drosophila*, and *C. elegans* to ~90% of the remaining known metazoan microRNA sequences. While this will not include the microRNAs which have been observed to be species-specific, a good deal of information from other model organisms is made available using an existing platform. This advantage has been useful as well in confirming the existence of novel microRNAs in one species that are homologous to known microRNAs. For example, the illumination of a probe for hsa-miR-519a from a labeled rat microRNA sample suggests the presence of a rat homolog for miR-519a. While this is at best circumstantial evidence, a list of potential homologs to be tested further could readily be obtained by hybridization to the existing platforms. It would be short-sighted to limit interpretation of microRNA array data to the given probes from one species.

Once a probe sequence is available to the public, a standard BLAST search against all known microRNAs will effectively "re-annotate" that specific probe. For the NCode™ microRNA array platform, this has already been conducted. A web tool is available (http://cord.rutgers.edu/gal_generator/) into which a blank array map can be input, and any number of available species can be selected. Probe identifiers are indexed and compared to previous BLAST results across all known microRNAs from all available species, identifying those probes designed for human microRNAs that are, for example, exact matches to gorilla. Probes found to vary by one or two nucleotides from a perfect match for a given species are labeled as negative controls to determine hybridization specificity. The output is a custom array layout file specifically annotated for any given species or combination of species, allowing the use of a single array to detect most microRNAs in all species found in miRBase. Take advantage of these highly conserved sequences when designing experiments. The lack of availability of a specific species microRNA microarray should not hinder the use of commercially available microRNA arrays.

Shi_Ch08.indd 149 5/30/2008 5:59:56 PM

8.6 Common Data Analysis Workflows for MicroRNA Microarrays

Much thought has been put into analysis of gene expression data over the past several years. Increasing use of microarray technology has highlighted the need for robust and accurate workflows for dealing with massive amounts of gene expression data. Many novel algorithms have emerged to deal with multivariate microarray data. For the most part, microRNA expression data can be treated with exactly the same methods as mRNA data. qPCR and array platforms have had to change very little to adapt to these smaller molecules and therefore standard workflows continue to apply. However, several of the traditional normalization methods are based on assumptions that do not hold true for current microRNA expression data.

As with most data analysis, the appropriate workflow is the one that makes the most sense in the context of the specific biological question being asked. In most cases, multiple arrays are used in a single experiment. This requires scaling and/or normalization methods to make the arrays comparable and compensate for artifacts or effects between arrays. Readily available techniques include a list of model-fitting approaches. Most model-fitting algorithms such as locally-weighted linear regression (loess), spline fitting (gspline), or linear modeling assume that there are a relatively large number (>1,000) of detectable measurements upon which to base interpretations. Additionally, most of these normalization methods assume that the majority of measurements will remain unchanged across the majority of the arrays (or conditions).

For smaller microRNA array experiments these assumptions may not be met. Since there are currently slightly less than 1,000 human microRNAs known, it is quite possible that a smaller microRNA dataset will not have a minimum number of measurements to meet this assumption. Reported measurements of microRNA expression levels are quite dynamic as well. The majority of microRNAs are expressed in a highly tissue-specific manner. This again would violate the assumption that genes remain relatively unchanged across arrays. For these otherwise limited datasets, we propose that a more appropriate choice for a normalization technique is a non-parametric method such as quantile normalization. Speed and colleagues were the first to apply quantile normalization to microarray datasets [19] and this was done originally on single-channel Affymetrix™ arrays. The only assumption is that the distribution of gene abundances is *nearly* the same in all samples. This is true for low abundance genes, and to a fairly good approximation, for genes of moderate abundance, but does not necessarily hold true for the few high-abundance genes, whose typical levels vary noticeably from sample to sample. In this normalization scheme, a gene X channel matrix is constructed from the dataset using background-subtracted intensity values. The matrix is then sorted by column into "quantiles" and the mean intensity value is taken across each row. This mean value then replaces the value in the original matrix order effectively forcing an identical distribution across all of the arrays. This brute-force normalization is an effective, rank-based method for reducing the effects of array and dye bias by

re-scaling the entire dataset. This can be accomplished easily in the R environment (http://www.r-project.org) using portions of the limma [155] package contained within Bioconductor [44] (http://www.bioconductor.org) (Panel 8.1).

Once the dataset has been normalized and corrected, you must now examine your experimental question. In most cases this is the identification of differentially expressed microRNAs. There are a wide variety of analyses from which to choose. For a simple two parameter comparison, the standard Student's t-test is often appropriate. We recommend SAM (Significance Analysis of Microarrays; http://www-stat.stanford.edu/~tibs/SAM/), a widely-used test similar to a t-test but including an estimate of false discovery error and designed specifically for microarray data. Standard ANOVA methods can also be used to explore the variance across more than two conditions. Regardless of the statistical test chosen, P-values must be corrected for multiple testing. With the high number of statistical tests (one for each gene), the likelihood of satisfying the null hypothesis (all means are equal across conditions) by chance alone increases. To take this into consideration, we must correct for performing multiple tests. The preferred method in microarray studies is to control the false discovery rate (FDR), or the expected proportion of incorrectly rejected null hypotheses. Benjamini-Hochberg [14], Bonferroni [62], and Westfall-Young [152, 153] are three commonly used types of multiple testing corrections; each is available from within R. The result will be a list of significantly regulated microRNAs that have passed the stringency test for multiple testing.

While this workflow is appropriate for small microRNA datasets, a more powerful and suitable approach is available for robust, multivariate expression datasets. Linear models have tremendous power to describe data, but have only recently become popular for microarray data analysis. Modeling a dataset entails the construction of a linear equation that "describes" the data based on a series of predefined parameters. The goal then is to be able to re-create the dataset with a minimal amount of parameters, while accounting for random errors. This is best

Panel 8.1 Quantile normalization in R. Quantile normalization can be performed fairly easily within the R environment. Begin by starting a session in the directory containing a comma-separated file containing a "gene x array" of background subtracted values. The following code and comments will describe the workflow within R:

```
#Load required library
library('limma')
#read file "file.csv" into object 'raw'
raw < -read.csv("file.csv", header = T)
#generate a box-whiskers plot for raw data to #determine variability across arrays
boxplot(log2(raw), main = "Raw")
#convert to matrix and normalize quantiles
norm < -normalizeQuantiles(as.matrix(raw))
norm < -as.data.frame(norm)
#generate box-whiskers plot for normalized data to confirm normalization
boxplot(log2(norm), main = "Normalized")
#write normalized values to new file
write.csv(norm,file = "norm.csv")
quit()
```

described by Kerr and Churchill [75–78], who describe a "minimal model" for two-color microarray data analysis that incorporates array, dye, gene, and sample effects (Panel 8.2). In addition to these parameters, combinatorial effects are also incorporated into their model to describe spot effects (array X gene), labeling effects (dye X gene), and gene-specific sample effects (sample X gene). It is this last effect that is most important as it describes the differential expression of a given gene across each sample. Care should be taken to balance the design when laying out an experiment for linear model analysis. Spreading your samples across multiple independent arrays and labeling the replicates with alternating dyes will help to estimate the technical errors produced by hybridization of individual arrays and dye labeling effects. Poor experimental balance can result in confounded parameters that cannot be estimated. Once the linear model has been described and fit to the dataset, F-tests are conducted on a per-gene basis comparing the model with and without the 'sample X gene' effects. This determines whether or not this effect contributes to a significant portion of the observed intensity value for the given gene, given the

Panel 8.2 Linear modeling of microarray data. The linear model concept attempts to mathematically define a dataset based on a given set of defined parameters. In essence, the experimenter describes the characteristics of the samples that are most important to a specific experiment. The "design" file is used to outline the parameters in an experiment and may be as simple as a "condition" parameter categorizing a sample as control or experimental, or significantly more complex. RNA source, sample preparation, technician name, or date of assay, are a small fraction of the parameters that can be included in a larger experiment. Each can be tested to determine if there is a significant "effect" on the resulting dataset. Once the parameters have been defined and associated with particular samples, a "comparison" matrix is used to describe the comparisons of interest among the many parameters. This is a fairly straightforward process for single-channel array data, but can become significantly more involved when dealing with two-color array data. To address this class of microarray data, Kerr and Churchill [75, 76, 78] proposed a standard linear model that attempts to describe some common sources of error in two-color microarray experiments. The model:

$$Y_{ijkg} = \mu + A_i + D_j + V_k + G_g + AG_{ig} + DG_{jg} + VG_{kg} + \varepsilon_{ijkg}$$

Where:

Y_{ijkg} = the observed values from the array experiment for the *i*th array, *j*th dye, *k*th sample, and *g*th gene.

μ = the estimated mean of the dataset.

A_i = the effect of being on the *i*th array.

D_j = the effect of being on the *j*th dye.

V_k = the effect of being on the *k*th sample.

G_g = the effect of being on the *g*th gene.

AG_{ig} = the combinatorial effect describing the spotting effect.

DG_{jg} = the combinatorial effect describing the gene-specific labeling effect.

VG_{kg} = the combinatorial effect describing the effect of being a given gene in a specific sample.

ε_{ijkg} = random error.

This model will take each of these parameters into consideration when fit to microarray data. The effect that is usually the most interesting is the VG_{kg} effect. This model can be fit with or without this specific effect and a per-gene F-test between the two fits will identify any genes with a considerable variance across any of the samples. P-values can be either tabulated from an F-distribution or, more appropriately, bootstrapped by finding the probability that a randomly permuted dataset will produce F-values greater than that observed with the original dataset. These P-values should be subject to the same multiple testing corrections used for standard array analysis.

distribution of the existing data. At this point, p-values can be bootstrapped by randomly permuting the dataset and conducting the same F-test. The probability of randomly obtaining a higher F-value will be determined empirically. P-values are adjusted for multiple comparisons using a standard 5% FDR cutoff.

The advantages of this data analysis approach are that few assumptions about the data are required *a priori*. The data themselves are used to drive the analysis. Ideally, linear modeling can be used to examine any number of parameters simultaneously (i.e., cell line, phenotype, differentiation state, etc.), and the more parameters that are included, the more accurate the resulting estimates of effect. The primary drawback is the rather large number of samples required to maintain sufficient degrees of freedom required for the estimation of each parameter. For each estimated value, a degree of freedom must be sacrificed. Since there are a large number of parameters involved (Array, Dye, Sample, Gene, etc.) this requires a robust dataset to make these estimates possible. The more replicates that are conducted, the more degrees of freedom available, yet experiment cost is a very common limiting factor inhibiting large numbers of replicates. This balance must be considered prior to accepting an experimental design, as the benefit of using linear modeling of microarray data is lost when you have to begin to sacrifice parameters to measure due to insufficient degrees of freedom. Linear modeling can be conducted using one of several freely available packages from the R/Bioconductor [44] environment. Limma [155] is a flexible and powerful linear modeling package that has extensive documentation, and is easily adaptable for both two-color as well as single-channel arrays. There is an accompanying graphical user interface (limmaGUI) that makes data input and analysis more streamlined and accessible. We routinely use the R/MAANOVA package [75, 76, 78] to analyze our microRNA data. This package provides detailed instructions for preparing your dataset and outlining the parameters for your experiment. In addition, MAANOVA provides a host of post-modeling features that help to streamline the analysis and interpretation of your significant genes.

8.7 Biclustering

The expression profiling of microRNAs is an important step in understanding the roles these molecules may play in regulating different cellular processes. However, what little we know about microRNAs suggests that these RNAs act exclusively through regulation of other genes. With this in mind, it then becomes prudent to examine the expression of the targets of these microRNAs under the same conditions, so as to (A) identify the pool of available targets under a given condition and (B) identify/predict specific targets for a subset of regulated microRNA based on the differential expression of target mRNAs. By directly comparing microRNA expression data to data obtained from other assays such as mRNA profiles, additional information can be gleaned as well. mRNAs known to have a role in transcriptional regulation have been shown to exert pressures on the promoter regions

of intergenic microRNAs (LAG, 2007, unpublished results). These interactions, as well as those satisfying other hypotheses, begin to emerge as underlying patterns in this combinatorial dataset.

The concept of biclustering has been used frequently to reveal relationships between genes based on their associations across numerous treatments or conditions [94, 103, 124, 147]. Logically, the tighter the association across multiple stressors, knockouts, treatments, or other conditions, the more confidence one would have that two genes are associated with similar biological processes and/or networks. The simplest approach to a biclustering analysis would be a cross-correlation study (Fig. 8.1). Basic hierarchical clustering across both genes and conditions would begin to unravel these relationships. This approach has been adapted by several groups to begin to reveal networks of relationships between microRNAs and their target mRNAs. One begins by constructing a matrix of correlations for all possible microRNA:mRNA pairs within a dataset. The cross-correlation, or correlation across these correlations, is then determined and used to relate neighboring microRNAs to each other based solely on their relationships across all observed mRNAs. The inverse is applied to mRNAs as well resulting in a two-dimensional hierarchical clustering matrix describing the relationships both within and between molecule types. We previously used cross-correlation clustering to investigate

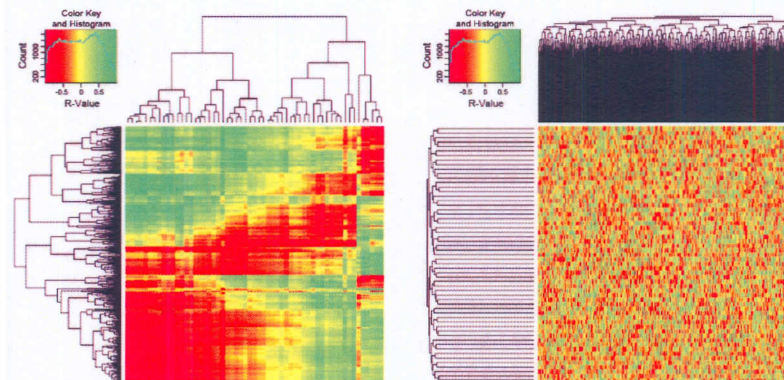

Fig. 8.1 The cross correlation matrix. Using a heatmap-like visualization technique, the correlations between microRNAs and mRNAs can be easily and readily visualized. In this experiment comparing human embryonic stem cells with embryoid bodies or embryonic carcinoma cells, regulated microRNAs (columns) are correlated to significant mRNAs (rows), and the resulting R-values are used to color the component blocks. Hierarchical clustering in both dimensions groups molecules based on their correlation across the opposing molecule type. Subclusters that are negatively correlated (red) may represent potential microRNA:mRNA target interactions resulting in degradation of the target mRNA. Regions showing strong positive correlation (green) may help to identify microRNAs and mRNAs that may have a shared functional pathway or transcriptional regulatory mechanism. The value of the biological information is seen when the connections between microRNA and mRNA are severed by randomly permuting the cross correlation matrix using the identical dataset, resulting in the complete abolition of subclusters

mRNAs and microRNAs differentially expressed between human embryonic stem cells, embryoid bodies, and embryonic carcinoma.

By focusing on subclusters of the cross-correlation matrix, several different interactions between microRNAs and mRNAs can be inferred. Since microRNAs have been shown to mediate mRNA degradation in certain conditions, it is reasonable to suggest that this activity could be identified in a subset, or bicluster, of the cross-correlation matrix demonstrating strong negative correlations across multiple conditions (actually, mRNA targeting would represent only one of two possible models of negative correlation). More specifically, if we can focus on a bicluster that contains mRNA that are downregulated as corresponding microRNA increase in abundance, we find the ideal subcluster to test putative interactions. A quick scan of a now more limited subset of mRNA may reveal a conserved microRNA binding motif amongst the candidate mRNA 3′UTRs. This approach is logical when we realize that our focus is on those mRNAs that are actively degraded as a result of microRNA binding activity. The extent of this mechanism in mammals is believed to be relatively little compared to translational repression or sequestration. The majority of microRNA:mRNA targets in mammals then, would probably not be identified in this manner. Interestingly, in apparent contrast to this challenge, a few groups have demonstrated an abundance of regulatory motifs in the 3′UTR of microRNAs downregulated during treatment with a specific microRNA [97]. The regulatory sequence contains a region of complementarity to the microRNA used. It may be that with the overexpression of a microRNA there is a detectable increase in the amount of target degradation that occurs. This suggests that targets can in fact be identified through their negative correlation to an enriched microRNA, but perhaps only in the context of extra-ordinary expression levels.

8.8 Revealing Regulation

While little is known about the functions of expressed microRNAs, even less is known about the mechanisms governing the regulation of microRNAs themselves. MicroRNAs are derived from both spliced intronic regions of mRNAs [98, 165, 166], as well as from unique transcripts located in intergenic regions [54, 88, 167]. Since no specific regulatory mechanism for microRNA processing has yet been identified, it is assumed that the expression of intronic microRNAs is regulated by the same mechanisms that regulate the abundance of the host transcript, as well as mechanisms governing intron splicing. Intergenic microRNAs have been shown to be transcribed by either Pol II [138] or in some cases by Pol III [20]. The Pol II transcripts are 5′-capped and poly(A)-tailed [22, 138] in a similar fashion to known mRNAs. A large number of intergenic microRNAs exist as poly-cistronic clusters. These clusters are often transcribed as a single unit and summarily processed into individual microRNA precursors after nuclear export [2, 59, 61, 65, 83, 138, 143]. Interestingly, early sequence analysis of upstream regions of intergenic microRNAs failed to identify common Pol II minimal promoter elements or similarities to

known mRNA promoter elements [22, 66, 138]. A few groups, including our lab, have since cloned and validated promoter regions for a small number of microRNAs, and demonstrated that similar regulatory mechanisms do in fact exist to control the transcription rates of intergenic microRNA. These mechanisms include conserved and occupied transcription factor binding sites as well as chromatin modifications, both of which have been shown to exert a regulatory pressure on the cloned promoters [41, 122, 131, 133, 141]. The inverse of the microRNA target analysis of a bicluster would be the determination of common regulatory mechanisms governing the expression levels of both microRNAs and mRNAs. Those transcripts sharing similar transcriptional regulatory mechanisms may be identified as microRNA: mRNA pairs demonstrating a positive correlation, and in most cases appropriate correlations with transcription factors that may be acting on these networks (i.e., positive correlation with activating transcription factors and negative correlation with transcriptional repressors).

Another approach that can be adopted to help unravel the regulatory mechanisms of microRNAs involves that bioinformatic prediction of a response to a specific transcription factor or factors based on upstream sequence analysis. We have recently applied this approach to study a specific pathway in differentiating mesenchymal stem cells (MSC) [49]. Upon identification of a specific pathway inhibitor (in this case, Tyrphostin AG-370) that significantly alters the ability of MSC to differentiate into osteocytes [120], we hypothesized that this effect may be mediated in part by regulated microRNA activity. Since AG-370 specifically inhibits the PDGF pathway, we conducted a literature search to identify transcription factors directly downstream of the PDGF receptors. Position weight matrices for each of these transcription factors were obtained from the TRANSFAC database [106, 158] and used to scan the 5 Kb upstream sequences of microRNAs that had been previously identified as significantly regulated during osteogenic differentiation of human MSCs. A comparison of the number of positive hits among the significant microRNAs to the number obtained from a sample of expressed but not regulated microRNA upstream sequences confirmed that these microRNAs were enriched for putative PDGF pathway binding sites ($p < 0.05$). A z-score analysis of individual microRNA upstream sequence hits vs. the average number of hits for all regulated and non-regulated upstream sequences was unable to identify individual microRNAs significantly enriched for binding sites after a multiple testing correction. However, the resulting p-value-rank-ordered list provided a confidence list that allowed us to rank microRNA regulatory sequences based on their predicted responses. Subsequent testing of regulated microRNA activity via qPCR during osteogenic differentiation in the presence of the inhibitor demonstrated that those microRNAs ranking higher in our confidence list showed a greater likelihood of being modulated by AG-370 treatment [49]. This workflow enables the prediction of microRNAs that may be responsive to a particular pathway or treatment, and the validation of these predictions via drug-targeting. This moderately high-throughput assay is just one example of a workflow that can be readily adapted to a wide variety of experimental questions pertinent to both microRNAs and stem cells.

8.9 Future Technologies

Current trends in technology will enable a fuller understanding of both the number and extent of microRNAs regulating stem cell function. Ultra high-throughput, deep sequencing technologies are beginning to emerge and re-define how nucleic acid sequences are identified, quantified, and regulated. Three competing yet similar technologies have emerged as the fore-runners in the field of deep sequencing. Each requires the preparation of a cDNA library, sequence amplification by PCR, high-density display of amplified sequences, and direct sequencing by either synthesis [15] or ligation methods [27]. With the ability to read, in parallel, upwards of 40 million ~35–50 bp sequences, the challenges of complexity and size of genetic information are readily addressed and ultra high-throughput assays become feasible for most investigators [10, 15, 69, 108, 128]. The application of these technologies to small RNAs is not lost, and in fact represents some of the first uses for deep sequencing. Direct sequencing of small RNAs requires no *a priori* knowledge of the microRNA sequence, which provides an immediate advantage for discovery over microarray technologies requiring the construction of complementary probes. However, access to a nearly-complete genome sequence is an important requirement since, in our experience, only ~50% of deep-sequencing microRNA tags can be aligned to genome, limiting the collection of valid data (LAG, 2008, unpublished results). However, by counting the frequencies of each unique ~35mer sequenced and validated by genome alignment, one can begin to examine the expression of specific sequences without the common complications of cross-hybridization, dye bias, microarray sensitivity, or saturation that plague microarray analyses. Counted data for each unique read can be interpreted as a direct measure of expression for use in differential expression studies.

As we begin to delve deeper into the genome in search of microRNAs and their targets, stem cells become an important piece of the puzzle. The previously described requirement of Dicer, DGCR8, and other members of the microRNA biogenesis pathway for both differentiation and maintenance of stem cells, combined with bioinformatic predictions of 10,000 microRNA genes, suggest that unknown numbers of novel sequences remain to be discovered in stem cells. A better appreciation of the mechanisms of stem cell regulation and the players involved will advance our understanding of the these crucial cells, and foster innovations in therapeutic applications of stem cells as well. The use of these next-generation techniques, however, must be accompanied by rigorous and novel statistical interpretations, as well as accommodations for the dimensions of the resulting data. Few biologists have had to deal with the volumes of data that will be generated in the near future, and even fewer have experience with mining and interpreting such large datasets. Collaborations with statisticians and computer scientists, and a focus on developing the next generation of biologists with strong working knowledge of both computer programming and statistical interpretation will be essential in discovering and interpreting microRNA regulatory mechanisms in differentiating stem cells.

Acknowledgements We thank Brad Love (Invitrogen) and Rebecka Jörnsten (Department of Statistics, Rutgers University) for patiently instructing us in statistical concepts and methods. LAG and RPH were supported by grants from NIH, the New Jersey Commission on Spinal Cord Research, the New Jersey Commission on Science and Technology, and Invitrogen, Inc.

References

1. Akao, Y., Y. Nakagawa, and T. Naoe, MicroRNAs 143 and 145 are possible common onco-microRNAs in human cancers. Oncol Rep, 2006. **16**(4): 845–50.
2. Altuvia, Y., P. Landgraf, G. Lithwick, N. Elefant, S. Pfeffer, A. Aravin, M.J. Brownstein, T. Tuschl, and H. Margalit, Clustering and conservation patterns of human microRNAs. Nucleic Acids Res, 2005. **33**(8): 2697–706.
3. Ambros, V., microRNAs: tiny regulators with great potential. Cell, 2001. **107**(7): 823–6.
4. Anderson, C., H. Catoe, and R. Werner, MIR-206 regulates connexin43 expression during skeletal muscle development. Nucleic Acids Res, 2006. **34**(20): 5863–71.
5. Babak, T., W. Zhang, Q. Morris, B.J. Blencowe, and T.R. Hughes, Probing microRNAs with microarrays: tissue specificity and functional inference. RNA, 2004. **10**(11): 1813.
6. Babak, T., W. Zhang, Q. Morris, B.J. Blencowe, and T.R. Hughes, Probing microRNAs with microarrays: tissue specificity and functional inference. RNA, 2004. **10**(11): 1813–9.
7. Baehrecke, E.H., miRNAs: micro managers of programmed cell death. Curr Biol, 2003. **13**(12): R473–5.
8. Barad, O., E. Meiri, A. Avniel, R. Aharonov, A. Barzilai, I. Bentwich, U. Einav, S. Gilad, P. Hurban, Y. Karov, E.K. Lobenhofer, E. Sharon, Y.M. Shiboleth, M. Shtutman, Z. Bentwich, and P. Einat, MicroRNA expression detected by oligonucleotide microarrays: system establishment and expression profiling in human tissues. Genome Res, 2004. **14**(12): 2486–94.
9. Baroukh, N., M.A. Ravier, M.K. Loder, E.V. Hill, A. Bounacer, R. Scharfmann, G.A. Rutter, and E. Van Obberghen, MicroRNA-124a regulates Foxa2 expression and intracellular signaling in pancreatic beta-cell lines. J Biol Chem, 2007. **282**(27): 19575–88.
10. Barski, A., S. Cuddapah, K. Cui, T.Y. Roh, D.E. Schones, Z. Wang, G. Wei, I. Chepelev, and K. Zhao, High-resolution profiling of histone methylations in the human genome. Cell, 2007. **129**(4): 823–37.
11. Bartel, D.P., MicroRNAs: genomics, biogenesis, mechanism, and function. Cell, 2004. **116**(2): 281–97.
12. Baskerville, S. and D.P. Bartel, Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. RNA, 2005. **11**(3): 241–7.
13. Behm-Ansmant, I., J. Rehwinkel, T. Doerks, A. Stark, P. Bork, and E. Izaurralde, mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1: DCP2 decapping complexes. Genes Dev, 2006. **20**(14): 1885–98.
14. Benjamini, Y. and Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc B, 1995. **57**: 289–300.
15. Bentley, D.R., Whole-genome re-sequencing. Curr Opin Genet Dev, 2006. **16**(6): 545–52.
16. Bentwich, I., A postulated role for microRNA in cellular differentiation. Faseb J, 2005. **19**(8): 875–9.
17. Bernstein, E., S.Y. Kim, M.A. Carmell, E.P. Murchison, H. Alcorn, M.Z. Li, A.A. Mills, S.J. Elledge, K.V. Anderson, and G.J. Hannon, Dicer is essential for mouse development. Nat Genet, 2003. **35**(3): 215–7.
18. Beuvink, I., F.A. Kolb, W. Budach, A. Garnier, J. Lange, F. Natt, U. Dengler, J. Hall, W. Filipowicz, and J. Weiler, A novel microarray approach reveals new tissue-specific signatures of known and predicted mammalian microRNAs. Nucleic Acids Res, 2007. **35**(7): e52.
19. Bolstad, B.M., R.A. Irizarry, M. Astrand, and T.P. Speed, A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics, 2003. **19**(2): 185–93.

20. Borchert, G.M., W. Lanier, and B.L. Davidson, RNA polymerase III transcribes human microRNAs. Nat Struct Mol Biol, 2006. **13**(12): 1097–101.
21. Brennecke, J., A. Stark, and S.M. Cohen, Not miR-ly muscular: microRNAs and muscle development. Genes Dev, 2005. **19**(19): 2261–4.
22. Cai, X., C.H. Hagedorn, and B.R. Cullen, Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA, 2004. **10**(12): 1957–66.
23. Calin, G.A., C.D. Dumitru, M. Shimizu, R. Bichi, S. Zupo, E. Noch, H. Aldler, S. Rattan, M. Keating, K. Rai, L. Rassenti, T. Kipps, M. Negrini, F. Bullrich, and C.M. Croce, Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A, 2002. **99**(24): 15524–9.
24. Calin, G.A., C.G. Liu, C. Sevignani, M. Ferracin, N. Felli, C.D. Dumitru, M. Shimizu, A. Cimmino, S. Zupo, M. Dono, M.L. Dell'Aquila, H. Alder, L. Rassenti, T.J. Kipps, F. Bullrich, M. Negrini, and C.M. Croce, MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. Proc Natl Acad Sci U S A, 2004. **101**(32): 11755.
25. Calin, G.A., C. Sevignani, C.D. Dumitru, T. Hyslop, E. Noch, S. Yendamuri, M. Shimizu, S. Rattan, F. Bullrich, M. Negrini, and C.M. Croce, Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A, 2004. **101**(9): 2999–3004.
26. Calin, G.A., M. Ferracin, A. Cimmino, G. Di Leva, M. Shimizu, S.E. Wojcik, M.V. Iorio, R. Visone, N.I. Sever, M. Fabbri, R. Iuliano, T. Palumbo, F. Pichiorri, C. Roldo, R. Garzon, C. Sevignani, L. Rassenti, H. Alder, S. Volinia, C.G. Liu, T.J. Kipps, M. Negrini, and C.M. Croce, A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med, 2005. **353**(17): 1793–801.
27. Carninci, P., T. Kasukawa, S. Katayama, J. Gough, M.C. Frith, N. Maeda, R. Oyama, T. Ravasi, B. Lenhard, C. Wells, R. Kodzius, K. Shimokawa, V.B. Bajic, S.E. Brenner, S. Batalov, A.R. Forrest, M. Zavolan, M.J. Davis, L.G. Wilming, V. Aidinis, J.E. Allen, A. Ambesi-Impiombato, R. Apweiler, R.N. Aturaliya, T.L. Bailey, M. Bansal, L. Baxter, K.W. Beisel, T. Bersano, H. Bono, A.M. Chalk, K.P. Chiu, V. Choudhary, A. Christoffels, D.R. Clutterbuck, M.L. Crowe, E. Dalla, B.P. Dalrymple, B. de Bono, G. Della Gatta, D. di Bernardo, T. Down, P. Engstrom, M. Fagiolini, G. Faulkner, C.F. Fletcher, T. Fukushima, M. Furuno, S. Futaki, M. Gariboldi, P. Georgii-Hemming, T.R. Gingeras, T. Gojobori, R.E. Green, S. Gustincich, M. Harbers, Y. Hayashi, T.K. Hensch, N. Hirokawa, D. Hill, L. Huminiecki, M. Iacono, K. Ikeo, A. Iwama, T. Ishikawa, M. Jakt, A. Kanapin, M. Katoh, Y. Kawasawa, J. Kelso, H. Kitamura, H. Kitano, G. Kollias, S.P. Krishnan, A. Kruger, S.K. Kummerfeld, I.V. Kurochkin, L.F. Lareau, D. Lazarevic, L. Lipovich, J. Liu, S. Liuni, S. McWilliam, M. Madan Babu, M. Madera, L. Marchionni, H. Matsuda, S. Matsuzawa, H. Miki, F. Mignone, S. Miyake, K. Morris, S. Mottagui-Tabar, N. Mulder, N. Nakano, H. Nakauchi, P. Ng, R. Nilsson, S. Nishiguchi, S. Nishikawa, F. Nori, O. Ohara, Y. Okazaki, V. Orlando, K.C. Pang, W.J. Pavan, G. Pavesi, G. Pesole, N. Petrovsky, S. Piazza, J. Reed, J.F. Reid, B.Z. Ring, M. Ringwald, B. Rost, Y. Ruan, S.L. Salzberg, A. Sandelin, C. Schneider, C. Schonbach, K. Sekiguchi, C.A. Semple, S. Seno, L. Sessa, Y. Sheng, Y. Shibata, H. Shimada, K. Shimada, D. Silva, B. Sinclair, S. Sperling, E. Stupka, K. Sugiura, R. Sultana, Y. Takenaka, K. Taki, K. Tammoja, S.L. Tan, S. Tang, M.S. Taylor, J. Tegner, S.A. Teichmann, H.R. Ueda, E. van Nimwegen, R. Verardo, C.L. Wei, K. Yagi, H. Yamanishi, E. Zabarovsky, S. Zhu, A. Zimmer, W. Hide, C. Bult, S.M. Grimmond, R.D. Teasdale, E.T. Liu, V. Brusic, J. Quackenbush, C. Wahlestedt, J.S. Mattick, D.A. Hume, C. Kai, D. Sasaki, Y. Tomaru, S. Fukuda, M. Kanamori-Katayama, M. Suzuki, J. Aoki, T. Arakawa, J. Iida, K. Imamura, M. Itoh, T. Kato, H. Kawaji, N. Kawagashira, T. Kawashima, M. Kojima, S. Kondo, H. Konno, K. Nakano, N. Ninomiya, T. Nishio, M. Okada, C. Plessy, K. Shibata, T. Shiraki, S. Suzuki, M. Tagami, K. Waki, A. Watahiki, Y. Okamura-Oho, H. Suzuki, J. Kawai and Y. Hayashizaki, The transcriptional landscape of the mammalian genome. Science, 2005. **309**(5740): 1559–63.
28. Castoldi, M., S. Schmidt, V. Benes, M. Noerholm, A.E. Kulozik, M.W. Hentze, and M.U. Muckenthaler, A sensitive array for microRNA expression profiling (miChip) based on locked nucleic acids (LNA). RNA, 2006. **12**(5): 913–20.

29. Chan, J.A., A.M. Krichevsky, and K.S. Kosik, MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. Cancer Res, 2005. **65**(14): 6029–33.
30. Chen, C., D.A. Ridzon, A.J. Broomer, Z. Zhou, D.H. Lee, J.T. Nguyen, M. Barbisin, N.L. Xu, V.R. Mahuvakar, M.R. Andersen, K.Q. Lao, K.J. Livak, and K.J. Guegler, Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res, 2005. **33**(20): e179.
31. Chen, C.Z. and H.F. Lodish, MicroRNAs as regulators of mammalian hematopoiesis. Semin Immunol, 2005. **17**(2): 155–65.
32. Chen, J.F., E.M. Mandel, J.M. Thomson, Q. Wu, T.E. Callis, S.M. Hammond, F.L. Conlon, and D.Z. Wang, The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. Nat Genet, 2006. **38**(2): 228–33.
33. Chen, X.M., P.L. Splinter, P. O'Hara S, and N.F. Larusso, A cellular micro-RNA, let-7i, regulates toll-like receptor 4 expression and contributes to cholangiocyte immune responses against cryptosporidium parvum Infection. J Biol Chem, 2007. **282**(39): 28929–38.
34. Cimmino, A., G.A. Calin, M. Fabbri, M.V. Iorio, M. Ferracin, M. Shimizu, S.E. Wojcik, R.I. Aqeilan, S. Zupo, M. Dono, L. Rassenti, H. Alder, S. Volinia, C.G. Liu, T.J. Kipps, M. Negrini, and C.M. Croce, miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci U S A, 2005. **102**(39): 13944–9.
35. Cullen, B.R., Derivation and function of small interfering RNAs and microRNAs. Virus Res, 2004. **102**(1): 3–9.
36. Davison, T.S., C.D. Johnson, and B.F. Andruss, Analyzing micro-RNA expression using microarrays. Methods Enzymol, 2006. **411**: 14–34.
37. Denli, A.M., B.B. Tops, R.H. Plasterk, R.F. Ketting, and G.J. Hannon, Processing of primary microRNAs by the Microprocessor complex. Nature, 2004. **432**(7014): 231–5.
38. Eder, M. and M. Scherr, MicroRNA and lung cancer. N Engl J Med, 2005. **352**(23): 2446–8.
39. Esau, C., X. Kang, E. Peralta, E. Hanson, E.G. Marcusson, L.V. Ravichandran, Y. Sun, S. Koo, R.J. Perera, R. Jain, N.M. Dean, S.M. Freier, C.F. Bennett, B. Lollo, and R. Griffey, MicroRNA-143 regulates adipocyte differentiation. J Biol Chem, 2004. **279**(50): 52361–5.
40. Esau, C., S. Davis, S.F. Murray, X.X. Yu, S.K. Pandey, M. Pear, L. Watts, S.L. Booten, M. Graham, R. McKay, A. Subramaniam, S. Propp, B.A. Lollo, S. Freier, C.F. Bennett, S. Bhanot, and B.P. Monia, miR-122 regulation of lipid metabolism revealed by in vivo antisense targeting. Cell Metab, 2006. **3**(2): 87–98.
41. Fazi, F., A. Rosa, A. Fatica, V. Gelmetti, M.L. De Marchis, C. Nervi, and I. Bozzoni, A minicircuitry comprised of microRNA-223 and transcription factors NFI-A and C/EBPalpha regulates human granulopoiesis. Cell, 2005. **123**(5): 819–31.
42. Felli, N., L. Fontana, E. Pelosi, R. Botta, D. Bonci, F. Facchiano, F. Liuzzi, V. Lulli, O. Morsilli, S. Santoro, M. Valtieri, G.A. Calin, C.G. Liu, A. Sorrentino, C.M. Croce, and C. Peschle, MicroRNAs 221 and 222 inhibit normal erythropoiesis and erythroleukemic cell growth via kit receptor down-modulation. Proc Natl Acad Sci U S A, 2005. **102**(50): 18081–6.
43. Frederikse, P.H., R. Donnelly, and L.M. Partyka, miRNA and Dicer in the mammalian lens: expression of brain-specific miRNAs in the lens. Histochem Cell Biol, 2006. **126**(1): 1–8.
44. Gentleman, R.C., V.J. Carey, D.M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A.J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J.Y. Yang, and J. Zhang, Bioconductor: open software development for computational biology and bioinformatics. Genome Biol, 2004. **5**(10): R80.
45. Giraldez, A.J., R.M. Cinalli, M.E. Glasner, A.J. Enright, J.M. Thomson, S. Baskerville, S.M. Hammond, D.P. Bartel, and A.F. Schier, MicroRNAs regulate brain morphogenesis in zebrafish. Science, 2005. **308**(5723): 833–8.
46. Giraldez, A.J., Y. Mishima, J. Rihel, R.J. Grocock, S. Van Dongen, K. Inoue, A.J. Enright, and A.F. Schier, Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science, 2006. **312**(5770): 75–9.

Shi_Ch08.indd 160 5/30/2008 5:59:59 PM

47. Goff, L.A., M. Yang, J. Bowers, R.C. Getts, R.W. Padgett, and R.P. Hart, Rational probe optimization and enhanced detection strategy for microRNAs using microarrays. RNA Biology, 2005. **2**(3): e9–e16.
48. Goff, L.A., M. Yang, J. Bowers, R.C. Getts, R.W. Padgett, and R.P. Hart, Rational probe optimization and enhanced detection strategy for microRNAs using microarrays. RNA Biol, 2005. **2**(3): 93–100.
[Au1] 49. Goff, L.A., S. Boucher, C. Ricupero, S. Fenstermacher, M. Swerdel, L. Chase, C. Adams, J.D. Chesnut, U. Lakshmipathy, and R.P. Hart, Differentiating human multipotent mesenchymal stromal cells regulate microRNAs: prediction of microRNA regulation by PDGF during osteogenesis (submitted, 2008).
50. Goodrich, J.A. and J.F. Kugel, Non-coding-RNA regulators of RNA polymerase II transcription. Nat Rev Mol Cell Biol, 2006. **7**(8): 612–6.
51. Gregory, R.I., K.P. Yan, G. Amuthan, T. Chendrimada, B. Doratotaj, N. Cooch, and R. Shiekhattar, The Microprocessor complex mediates the genesis of microRNAs. Nature, 2004. **432**(7014): 235–40.
52. Gregory, R.I. and R. Shiekhattar, MicroRNA biogenesis and cancer. Cancer Res, 2005. **65**(9): 3509–12.
53. Griffiths-Jones, S., R.J. Grocock, S. van Dongen, A. Bateman, and A.J. Enright, miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res, 2006. **34**(Database issue): D140–4.
54. Gu, J., T. He, Y. Pei, F. Li, X. Wang, J. Zhang, X. Zhang, and Y. Li, Primary transcripts and expressions of mammal intergenic microRNAs detected by mapping ESTs to their flanking sequences. Mamm Genome, 2006. **17**(10): 1033–41.
55. Han, J., Y. Lee, K.H. Yeom, Y.K. Kim, H. Jin, and V.N. Kim, The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev, 2004. **18**(24): 3016–27.
56. Han, J., Y. Lee, K.H. Yeom, J.W. Nam, I. Heo, J.K. Rhee, S.Y. Sohn, Y. Cho, B.T. Zhang, and V.N. Kim, Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell, 2006. **125**(5): 887–901.
57. Harfe, B.D., M.T. McManus, J.H. Mansfield, E. Hornstein, and C.J. Tabin, The RNaseIII enzyme Dicer is required for morphogenesis but not patterning of the vertebrate limb. Proc Natl Acad Sci U S A, 2005. **102**(31): 10898–903.
58. Hatfield, S.D., H.R. Shcherbata, K.A. Fischer, K. Nakahara, R.W. Carthew, and H. Ruohola-Baker, Stem cell division is regulated by the microRNA pathway. Nature, 2005. **435**(7044): 974–8.
59. Hayashita, Y., H. Osada, Y. Tatematsu, H. Yamada, K. Yanagisawa, S. Tomida, Y. Yatabe, K. Kawahara, Y. Sekido, and T. Takahashi, A polycistronic microRNA cluster, miR-17–92, is overexpressed in human lung cancers and enhances cell proliferation. Cancer Res, 2005. **65**(21): 9628–32.
60. He, H., K. Jazdzewski, W. Li, S. Liyanarachchi, R. Nagy, S. Volinia, G.A. Calin, C.G. Liu, K. Franssila, S. Suster, R.T. Kloos, C.M. Croce, and A. de la Chapelle, The role of microRNA genes in papillary thyroid carcinoma. Proc Natl Acad Sci U S A, 2005. **102**(52): 19075–80.
61. He, L., J.M. Thomson, M.T. Hemann, E. Hernando-Monge, D. Mu, S. Goodson, S. Powers, C. Cordon-Cardo, S.W. Lowe, G.J. Hannon, and S.M. Hammond, A microRNA polycistron as a potential human oncogene. Nature, 2005. **435**(7043): 828–33.
62. Holm, S., A simple sequentially rejective bonferroni test procedure. Scand J Stat, 1979. **6**: 65–70.
63. Hornstein, E., J.H. Mansfield, S. Yekta, J.K. Hu, B.D. Harfe, M.T. McManus, S. Baskerville, D.P. Bartel, and C.J. Tabin, The microRNA miR-196 acts upstream of Hoxb8 and Shh in limb development. Nature, 2005. **438**(7068): 671–4.
64. Hossain, A., M.T. Kuo, and G.F. Saunders, Mir-17–5p regulates breast cancer cell proliferation by inhibiting translation of AIB1 mRNA. Mol Cell Biol, 2006. **26**(21): 8191–201.
65. Houbaviy, H.B., M.F. Murray, and P.A. Sharp, Embryonic stem cell-specific MicroRNAs. Dev Cell, 2003. **5**(2): 351–8.

66. Houbaviy, H.B., L. Dennis, R. Jaenisch, and P.A. Sharp, Characterization of a highly variable eutherian microRNA gene. RNA, 2005. **11**(8): 1245–57.
67. Huang, J., F. Wang, E. Argyris, K. Chen, Z. Liang, H. Tian, W. Huang, K. Squires, G. Verlinghieri, and H. Zhang, Cellular microRNAs contribute to HIV-1 latency in resting primary CD4(+) T lymphocytes. Nat Med, 2007. **13**(10): 1241–7.
68. Humphreys, D.T., B.J. Westman, D.I. Martin, and T. Preiss, MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. Proc Natl Acad Sci U S A, 2005. **102**(47): 16961–6.
69. Johnson, D.S., A. Mortazavi, R.M. Myers, and B. Wold, Genome-wide mapping of in vivo protein-DNA interactions. Science, 2007. **316**(5830): 1497–502.
70. Johnson, S.M., H. Grosshans, J. Shingara, M. Byrom, R. Jarvis, A. Cheng, E. Labourier, K.L. Reinert, D. Brown, and F.J. Slack, RAS is regulated by the let-7 microRNA family. Cell, 2005. **120**(5): 635–47.
71. Johnston, R.J. and O. Hobert, A microRNA controlling left/right neuronal asymmetry in Caenorhabditis elegans. Nature, 2003. **426**(6968): 845–9.
72. Josephson, R., C.J. Ording, Y. Liu, S. Shin, U. Lakshmipathy, A. Toumadje, B. Love, J.D. Chesnut, P.W. Andrews, M.S. Rao, and J.M. Auerbach, Qualification of embryonal carcinoma 2102Ep as a reference for human embryonic stem cell research. Stem Cells, 2007. **25**(2): 437–46.
73. Kanellopoulou, C., S.A. Muljo, A.L. Kung, S. Ganesan, R. Drapkin, T. Jenuwein, D.M. Livingston, and K. Rajewsky, Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev, 2005. **19**(4): 489–501.
74. Kawasaki, H. and K. Taira, Functional analysis of microRNAs during the retinoic acid-induced neuronal differentiation of human NT2 cells. Nucleic Acids Res Suppl, 2003. (3): 243–4.
75. Kerr, M.K., M. Martin, and G.A. Churchill, Analysis of variance for gene expression microarray data. J Comput Biol, 2000. **7**(6): 819–37.
76. Kerr, M.K. and G.A. Churchill, Experimental design for gene expression microarrays. Biostatistics, 2001. **2**(2): 183–201.
77. Kerr, M.K. and G.A. Churchill, Bootstrapping cluster analysis: assessing the reliability of conclusions from microarray experiments. Proc Natl Acad Sci U S A, 2001. **98**(16): 8961–5.
78. Kerr, M.K. and G.A. Churchill, Statistical design and the analysis of gene expression microarray data. Genet Res, 2001. **77**(2): 123–8.
79. Kim, H.K., Y.S. Lee, U. Sivaprasad, A. Malhotra, and A. Dutta, Muscle-specific microRNA miR-206 promotes muscle differentiation. J Cell Biol, 2006. **174**(5): 677–87.
80. Kim, J., A. Krichevsky, Y. Grad, G.D. Hayes, K.S. Kosik, G.M. Church, and G. Ruvkun, Identification of many microRNAs that copurify with polyribosomes in mammalian neurons. Proc Natl Acad Sci U S A, 2004. **101**(1): 360–5.
81. Kluiver, J., E. Haralambieva, D. de Jong, T. Blokzijl, S. Jacobs, B.J. Kroesen, S. Poppema, and A. van den Berg, Lack of BIC and microRNA miR-155 expression in primary cases of Burkitt lymphoma. Genes Chromosomes Cancer, 2006. **45**(2): 147–53.
82. Krichevsky, A.M., K.S. King, C.P. Donahue, K. Khrapko, and K.S. Kosik, A microRNA array reveals extensive regulation of microRNAs during brain development. RNA, 2003. **9**(10): 1274.
83. Krichevsky, A.M., K.S. King, C.P. Donahue, K. Khrapko, and K.S. Kosik, A microRNA array reveals extensive regulation of microRNAs during brain development. RNA, 2003. **9**(10): 1274–81.
84. Kutay, H., S. Bai, J. Datta, T. Motiwala, I. Pogribny, W. Frankel, S.T. Jacob, and K. Ghoshal, Downregulation of miR-122 in the rodent and human hepatocellular carcinomas. J Cell Biochem, 2006. **99**(3): 671–8.
85. Kuwabara, T., J. Hsieh, K. Nakashima, K. Taira, and F.H. Gage, A small modulatory dsRNA specifies the fate of adult neural stem cells. Cell, 2004. **116**(6): 779–93.

86. Kwon, C., Z. Han, E.N. Olson, and D. Srivastava, MicroRNA1 influences cardiac differentiation in Drosophila and regulates Notch signaling. Proc Natl Acad Sci U S A, 2005. **102**(52): 18986–91.
87. Kye, M.J., T. Liu, S.F. Levy, N.L. Xu, B.B. Groves, R. Bonneau, K. Lao, and K.S. Kosik, Somatodendritic microRNAs identified by laser capture and multiplex RT-PCR. RNA, 2007. **13**(8): 1224–34.
88. Lagos-Quintana, M., R. Rauhut, J. Meyer, A. Borkhardt, and T. Tuschl, New microRNAs from mouse and human. RNA, 2003. **9**(2): 175–9.
89. Lakshmipathy, U., B. Love, L.A. Goff, R. Jornsten, R. Graichen, R.P. Hart, and J.D. Chesnut, MicroRNA expression pattern of undifferentiated and differentiated human embryonic stem cells. Stem Cells Dev, 2007. **16**(6): 1003–16.
90. Lecellier, C.H., P. Dunoyer, K. Arar, J. Lehmann-Che, S. Eyquem, C. Himber, A. Saib, and O. Voinnet, A cellular microRNA mediates antiviral defense in human cells. Science, 2005. **308**(5721): 557–60.
91. Lee, Y., C. Ahn, J. Han, H. Choi, J. Kim, J. Yim, J. Lee, P. Provost, O. Radmark, S. Kim, and V.N. Kim, The nuclear RNase III Drosha initiates microRNA processing. Nature, 2003. **425**(6956): 415–9.
92. Lee, Y.S., H.K. Kim, S. Chung, K.S. Kim, and A. Dutta, Depletion of human micro-RNA miR-125b reveals that it is critical for the proliferation of differentiated cells but not for the down-regulation of putative targets during differentiation. J Biol Chem, 2005. **280**(17): 16635–41.
93. Lewis, B.P., C.B. Burge, and D.P. Bartel, Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell, 2005. **120**(1): 15–20.
94. Li, H., X. Chen, K. Zhang, and T. Jiang, A general framework for biclustering gene expression data. J Bioinform Comput Biol, 2006. **4**(4): 911–33.
95. Lian, S., A. Jakymiw, T. Eystathioy, J.C. Hamel, M.J. Fritzler, and E.K. Chan, GW bodies, microRNAs and the cell cycle. Cell Cycle, 2006. **5**(3): 242–5.
96. Liang, R.Q., W. Li, Y. Li, C.Y. Tan, J.X. Li, Y.X. Jin, and K.C. Ruan, An oligonucleotide microarray for microRNA expression analysis based on labeling RNA with quantum dot and nanogold probe. Nucleic Acids Res, 2005. **33**(2): e17.
97. Lim, L.P., N.C. Lau, P. Garrett-Engele, A. Grimson, J.M. Schelter, J. Castle, D.P. Bartel, P.S. Linsley, and J.M. Johnson, Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature, 2005. **433**(7027): 769–73.
98. Lin, S.L., J.D. Miller, and S.Y. Ying, Intronic MicroRNA (miRNA). J Biomed Biotechnol, 2006. **2006**(4): 26818.
99. Liu, C.G., G.A. Calin, B. Meloon, N. Gamliel, C. Sevignani, M. Ferracin, C.D. Dumitru, M. Shimizu, S. Zupo, M. Dono, H. Alder, F. Bullrich, M. Negrini, and C.M. Croce, An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. Proc Natl Acad Sci U S A, 2004. **101**(26): 9740.
100. Liu, J., M.A. Carmell, F.V. Rivas, C.G. Marsden, J.M. Thomson, J.J. Song, S.M. Hammond, L. Joshua-Tor, and G.J. Hannon, Argonaute2 is the catalytic engine of mammalian RNAi. Science, 2004. **305**(5689): 1437–41.
101. Liu, J., F.V. Rivas, J. Wohlschlegel, J.R. Yates, 3rd, R. Parker, and G.J. Hannon, A role for the P-body component GW182 in microRNA function. Nat Cell Biol, 2005. **7**(12): 1261–6.
102. Liu, J., M.A. Valencia-Sanchez, G.J. Hannon, and R. Parker, MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. Nat Cell Biol, 2005. **7**(7): 719–23.
103. Madeira, S.C. and A.L. Oliveira, Biclustering algorithms for biological data analysis: a survey. IEEE/ACM Trans Comput Biol Bioinform, 2004. **1**(1): 24–45.
104. Mansfield, J.H., B.D. Harfe, R. Nissen, J. Obenauer, J. Srineel, A. Chaudhuri, R. Farzan-Kashani, M. Zuker, A.E. Pasquinelli, G. Ruvkun, P.A. Sharp, C.J. Tabin, and M.T. McManus, MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. Nat Genet, 2004. **36**(10): 1079–83.

105. Masotti, A., Preckel, T., Analysis of small RNAs with the Agilent 2100 Bioanalyzer. Nature Methods, 2006. **3**(8): iii–iv.
106. Matys, V., E. Fricke, R. Geffers, E. Gossling, M. Haubrock, R. Hehl, K. Hornischer, D. Karas, A.E. Kel, O.V. Kel-Margoulis, D.U. Kloos, S. Land, B. Lewicki-Potapov, H. Michael, R. Munch, I. Reuter, S. Rotert, H. Saxel, M. Scheer, S. Thiele, and E. Wingender, TRANSFAC: transcriptional regulation, from patterns to profiles. Nucleic Acids Res, 2003. **31**(1): 374–8.
107. Matzke, M., W. Aufsatz, T. Kanno, L. Daxinger, I. Papp, M.F. Mette, and A.J. Matzke, Genetic analysis of RNA-mediated transcriptional gene silencing. Biochim Biophys Acta, 2004. **1677**(1–3): 129–41.
108. Mikkelsen, T.S., M. Ku, D.B. Jaffe, B. Issac, E. Lieberman, G. Giannoukos, P. Alvarez, W. Brockman, T.K. Kim, R.P. Koche, W. Lee, E. Mendenhall, A. O'Donovan, A. Presser, C. Russ, X. Xie, A. Meissner, M. Wernig, R. Jaenisch, C. Nusbaum, E.S. Lander, and B.E. Bernstein, Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature, 2007. **448**(7153): 553–60.
109. Miranda, K.C., T. Huynh, Y. Tay, Y.S. Ang, W.L. Tam, A.M. Thomson, B. Lim, and I. Rigoutsos, A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell, 2006. **126**(6): 1203–17.
110. Mishima, Y., A.J. Giraldez, Y. Takeda, T. Fujiwara, H. Sakamoto, A.F. Schier, and K. Inoue, Differential regulation of germline mRNAs in soma and germ cells by zebrafish miR-430. Curr Biol, 2006. **16**(21): 2135–42.
111. Miska, E.A., E. Alvarez-Saavedra, M. Townsend, A. Yoshii, N. Sestan, P. Rakic, M. Constantine-Paton, and H.R. Horvitz, Microarray analysis of microRNA expression in the developing mammalian brain. Genome Biol, 2004. **5**(9): R68.
112. Moffat, I.D., P.C. Boutros, T. Celius, J. Linden, R. Pohjanvirta, and A.B. Okey, MicroRNAs in adult rodent liver are refractory to dioxin treatment. Toxicol Sci, 2007. **99**(2): 470–87.
113. Mora, J.R. and R.C. Getts, Enzymatic microRNA detection in microtiter plates with DNA dendrimers. BioTechniques, 2006. **41**(4): 420, 422, 424.
114. Murchison, E.P., J.F. Partridge, O.H. Tam, S. Cheloufi, and G.J. Hannon, Characterization of Dicer-deficient murine embryonic stem cells. Proc Natl Acad Sci U S A, 2005. **102**(34): 12135–40.
115. Naguibneva, I., M. Ameyar-Zazoua, A. Polesskaya, S. Ait-Si-Ali, R. Groisman, M. Souidi, S. Cuvellier, and A. Harel-Bellan, The microRNA miR-181 targets the homeobox protein Hox-A11 during mammalian myoblast differentiation. Nat Cell Biol, 2006. **8**(3): 278–84.
116. Nakajima, N., T. Takahashi, R. Kitamura, K. Isodono, S. Asada, T. Ueyama, H. Matsubara, and H. Oh, MicroRNA-1 facilitates skeletal myogenic differentiation without affecting osteoblastic and adipogenic differentiation. Biochem Biophys Res Commun, 2006. **350**(4): 1006–12.
117. Nelson, P.T., D.A. Baldwin, L.M. Scearce, J.C. Oberholtzer, J.W. Tobias, and Z. Mourelatos, Microarray-based, high-throughput gene expression profiling of microRNAs. Nat Methods, 2004. **1**(2): 155–61.
118. Nelson, P.T., A.G. Hatzigeorgiou, and Z. Mourelatos, miRNP:mRNA association in polyribosomes in a human neuronal cell line. RNA, 2004. **10**(3): 387–94.
119. Ng, F., S. Boucher, S. Koh, K.S. Sastry, L. Chase, U. Lakshmipathy, C. Choong, Z. Yang, M. Vemuri, M.S. Rao, and V. Tanavde, PDGF, TGF-b and FGF signaling is important for differentiation and growth of Mesenchymal Stem Cells (MSCs): transcriptional profiling can identify markers and signaling pathways important in differentiation of MSC into adipogenic, chondrogenic and osteogenic lineages. Blood, 2008 Mar 10 (Epub ahead of print)
120. Pauley, K.M., T. Eystathioy, A. Jakymiw, J.C. Hamel, M.J. Fritzler, and E.K. Chan, Formation of GW bodies is a consequence of microRNA genesis. EMBO Rep, 2006. **7**(9): 904–10.
121. Pillai, R.S., S.N. Bhattacharyya, C.G. Artus, T. Zoller, N. Cougot, E. Basyuk, E. Bertrand, and W. Filipowicz, Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science, 2005. **309**(5740): 1573–6.

122. Plaisance, V., A. Abderrahmani, V. Perret-Menoud, P. Jacquemin, F. Lemaigre, and R. Regazzi, MicroRNA-9 controls the expression of Granuphilin/Slp4 and the secretory response of insulin-producing cells. J Biol Chem, 2006. **281**(37): 26932–42.
123. Poy, M.N., L. Eliasson, J. Krutzfeldt, S. Kuwajima, X. Ma, P.E. Macdonald, S. Pfeffer, T. Tuschl, N. Rajewsky, P. Rorsman, and M. Stoffel, A pancreatic islet-specific microRNA regulates insulin secretion. Nature, 2004. **432**(7014): 226–30.
124. Prelic, A., S. Bleuler, P. Zimmermann, A. Wille, P. Buhlmann, W. Gruissem, L. Hennig, L. Thiele, and E. Zitzler, A systematic comparison and evaluation of biclustering methods for gene expression data. Bioinformatics, 2006. **22**(9): 1122–9.
125. Ramkissoon, S.H., L.A. Mainwaring, Y. Ogasawara, K. Keyvanfar, J.P. McCoy, Jr., E.M. Sloand, S. Kajigaya, and N.S. Young, Hematopoietic-specific microRNA expression in human cells. Leuk Res, 2006. **30**(5): 643–7.
126. Rao, M., Conserved and divergent paths that regulate self-renewal in mouse and human embryonic stem cells. Dev Biol, 2004. **275**(2): 269–86.
127. Rehwinkel, J., I. Behm-Ansmant, D. Gatfield, and E. Izaurralde, A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. RNA, 2005. **11**(11): 1640–7.
128. Robertson, G., M. Hirst, M. Bainbridge, M. Bilenky, Y. Zhao, T. Zeng, G. Euskirchen, B. Bernier, R. Varhol, A. Delaney, N. Thiessen, O.L. Griffith, A. He, M. Marra, M. Snyder, and S. Jones, Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. Nat Methods, 2007. **4**(8): 651–7.
129. Roldo, C., E. Missiaglia, J.P. Hagan, M. Falconi, P. Capelli, S. Bersani, G.A. Calin, S. Volinia, C.G. Liu, A. Scarpa, and C.M. Croce, MicroRNA expression abnormalities in pancreatic endocrine and acinar tumors are associated with distinctive pathologic features and clinical behavior. J Clin Oncol, 2006. **24**(29): 4677–84.
130. Saito, K., K.M. Nishida, T. Mori, Y. Kawamura, K. Miyoshi, T. Nagami, H. Siomi, and M.C. Siomi, Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. Genes Dev, 2006. **20**(16): 2214–22.
131. Saito, Y., G. Liang, G. Egger, J.M. Friedman, J.C. Chuang, G.A. Coetzee, and P.A. Jones, Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. Cancer Cell, 2006. **9**(6): 435–43.
132. Schramke, V. and R. Allshire, Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin-based gene silencing. Science, 2003. **301**(5636): 1069–74.
133. Scott, G.K., M.D. Mattie, C.E. Berger, S.C. Benz, and C.C. Benz, Rapid alteration of microRNA levels by histone deacetylase inhibition. Cancer Res, 2006. **66**(3): 1277–81.
134. Sempere, L.F., S. Freemantle, I. Pitha-Rowe, E. Moss, E. Dmitrovsky, and V. Ambros, Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. Genome Biol, 2004. **5**(3): R13.
135. Shcherbata, H.R., S. Hatfield, E.J. Ward, S. Reynolds, K.A. Fischer, and H. Ruohola-Baker, The MicroRNA pathway plays a regulatory role in stem cell division. Cell Cycle, 2006. **5**(2): 172–5.
136. Shi, L., L.H. Reid, W.D. Jones, R. Shippy, J.A. Warrington, S.C. Baker, P.J. Collins, F. de Longueville, E.S. Kawasaki, K.Y. Lee, Y. Luo, Y.A. Sun, J.C. Willey, R.A. Setterquist, G.M. Fischer, W. Tong, Y.P. Dragan, D.J. Dix, F.W. Frueh, F.M. Goodsaid, D. Herman, R.V. Jensen, C.D. Johnson, E.K. Lobenhofer, R.K. Puri, U. Schrf, J. Thierry-Mieg, C. Wang, M. Wilson, P.K. Wolber, L. Zhang, S. Amur, W. Bao, C.C. Barbacioru, A.B. Lucas, V. Bertholet, C. Boysen, B. Bromley, D. Brown, A. Brunner, R. Canales, X.M. Cao, T.A. Cebula, J.J. Chen, J. Cheng, T.M. Chu, E. Chudin, J. Corson, J.C. Corton, L.J. Croner, C. Davies, T.S. Davison, G. Delenstarr, X. Deng, D. Dorris, A.C. Eklund, X.H. Fan, H. Fang, S. Fulmer-Smentek, J.C. Fuscoe, K. Gallagher, W. Ge, L. Guo, X. Guo, J. Hager, P.K. Haje, J. Han, T. Han, H.C. Harbottle, S.C. Harris, E. Hatchwell, C.A. Hauser, S. Hester, H. Hong, P. Hurban, S.A. Jackson, H. Ji, C.R. Knight, W.P. Kuo, J.E. LeClerc, S. Levy, Q.Z. Li, C. Liu, Y. Liu, M.J. Lombardi, Y. Ma, S.R. Magnuson, B. Maqsodi, T. McDaniel, N. Mei, O. Myklebost, B.

Shi_Ch08.indd 165 5/30/2008 6:00:00 PM

Ning, N. Novoradovskaya, M.S. Orr, T.W. Osborn, A. Papallo, T.A. Patterson, R.G. Perkins, E.H. Peters, R. Peterson, K.L. Philips, P.S. Pine, L. Pusztai, F. Qian, H. Ren, M. Rosen, B.A. Rosenzweig, R.R. Samaha, M. Schena, G.P. Schroth, S. Shchegrova, D.D. Smith, F. Staedtler, Z. Su, H. Sun, Z. Szallasi, Z. Tezak, D. Thierry-Mieg, K.L. Thompson, I. Tikhonova, Y. Turpaz, B. Vallanat, C. Van, S.J. Walker, S.J. Wang, Y. Wang, R. Wolfinger, A. Wong, J. Wu, C. Xiao, Q. Xie, J. Xu, W. Yang, S. Zhong, Y. Zong and W. Slikker, Jr., The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nat Biotechnol, 2006. **24**(9): 1151–61.

137. Song, L. and R.S. Tuan, MicroRNAs and cell differentiation in mammalian development. Birth Defects Res C Embryo Today, 2006. **78**(2): 140–9.
138. Suh, M.R., Y. Lee, J.Y. Kim, S.K. Kim, S.H. Moon, J.Y. Lee, K.Y. Cha, H.M. Chung, H.S. Yoon, S.Y. Moon, V.N. Kim, and K.S. Kim, Human embryonic stem cells express a unique set of microRNAs. Dev Biol, 2004. **270**(2): 488–98.
139. Sun, Y., S. Koo, N. White, E. Peralta, C. Esau, N.M. Dean, and R.J. Perera, Development of a micro-array to detect human and mouse microRNAs and characterization of expression in human organs. Nucleic Acids Res, 2004. **32**(22): e188.
140. Sweetman, D., T. Rathjen, M. Jefferson, G. Wheeler, T.G. Smith, G.N. Wheeler, A. Munsterberg, and T. Dalmay, FGF-4 signaling is involved in mir-206 expression in developing somites of chicken embryos. Dev Dyn, 2006. **235**(8): 2185–91.
141. Taganov, K.D., M.P. Boldin, K.J. Chang, and D. Baltimore, NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. Proc Natl Acad Sci U S A, 2006. **103**(33): 12481–6.
142. Tam, W. and J.E. Dahlberg, miR-155/BIC as an oncogenic microRNA. Genes Chromosomes Cancer, 2006. **45**(2): 211–2.
143. Tanzer, A. and P.F. Stadler, Molecular evolution of a microRNA cluster. J Mol Biol, 2004. **339**(2): 327–35.
144. Tay, Y.M., W.L. Tam, Y.S. Ang, P.M. Gaughwin, H. Yang, W. Wang, R. Liu, J. George, H.H. Ng, R.J. Perera, T. Lufkin, I. Rigoutsos, A.M. Thomson, and B. Lim, MicroRNA-134 modulates the differentiation of mouse embryonic stem cells, where it causes post-transcriptional attenuation of Nanog and LRH1. Stem Cells, 2008. **26**(1): 17–29.
145. Thomson, J.M., J. Parker, C.M. Perou, and S.M. Hammond, A custom microarray platform for analysis of microRNA gene expression. Nat Methods, 2004. **1**(1): 47–53.
146. Tsuchiya, Y., M. Nakajima, S. Takagi, T. Taniya, and T. Yokoi, MicroRNA regulates the expression of human cytochrome P450 1B1. Cancer Res, 2006. **66**(18): 9090–8.
147. Turner, H.L., T.C. Bailey, W.J. Krzanowski, and C.A. Hemingway, Biclustering models for structured microarray data. IEEE/ACM Trans Comput Biol Bioinform, 2005. **2**(4): 316–29.
148. van Rooij, E., L.B. Sutherland, N. Liu, A.H. Williams, J. McAnally, R.D. Gerard, J.A. Richardson, and E.N. Olson, A signature pattern of stress-responsive microRNAs that can evoke cardiac hypertrophy and heart failure. Proc Natl Acad Sci U S A, 2006. **103**(48): 18255–60.
149. Voorhoeve, P.M., C. le Sage, M. Schrier, A.J. Gillis, H. Stoop, R. Nagel, Y.P. Liu, J. van Duijse, J. Drost, A. Griekspoor, E. Zlotorynski, N. Yabuta, G. De Vita, H. Nojima, L.H. Looijenga, and R. Agami, A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors. Cell, 2006. **124**(6): 1169–81.
150. Wang, H., R.A. Ach, and B. Curry, Direct and sensitive miRNA profiling from low-input total RNA. RNA, 2007. **13**(1): 151–9.
151. Wang, Y., R. Medvid, C. Melton, R. Jaenisch, and R. Blelloch, DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet, 2007. **39**(3): 380–5.
152. Westfall, P.H. and S.S. Young, Resampling-Based Multiple Testing. 1993, New York: Wiley.
153. Westfall, P.H., D.V. Zaykin, and S.S. Young, Multiple tests for genetic effects in association studies. Methods Mol Biol, 2002. **184**: 143–68.

154. Weston, M.D., M.L. Pierce, S. Rocha-Sanchez, K.W. Beisel, and G.A. Soukup, MicroRNA gene expression in the mouse inner ear. Brain Res, 2006. **1111**(1): 95–104.
155. Wettenhall, J.M. and G.K. Smyth, limmaGUI: a graphical user interface for linear modeling of microarray data. Bioinformatics, 2004. **20**(18): 3705–6.
156. Wienholds, E., M.J. Koudijs, F.J. van Eeden, E. Cuppen, and R.H. Plasterk, The microRNA-producing enzyme Dicerl is essential for zebrafish development. Nat Genet, 2003. **35**(3): 217–8.
157. Wienholds, E., W.P. Kloosterman, E. Miska, E. Alvarez-Saavedra, E. Berezikov, E. de Bruijn, H.R. Horvitz, S. Kauppinen, and R.H. Plasterk, MicroRNA expression in zebrafish embryonic development. Science, 2005. **309**(5732): 310–1.
158. Wingender, E., P. Dietze, H. Karas, and R. Knuppel, TRANSFAC: a database on transcription factors and their DNA binding sites. Nucleic Acids Res, 1996. **24**(1): 238–41.
159. Wu, L. and J.G. Belasco, Micro-RNA regulation of the mammalian lin-28 gene during neuronal differentiation of embryonal carcinoma cells. Mol Cell Biol, 2005. **25**(21): 9198–208.
160. Wu, L., J. Fan, and J.G. Belasco, MicroRNAs direct rapid deadenylation of mRNA. Proc Natl Acad Sci U S A, 2006. **103**(11): 4034–9.
161. Xu, P., S.Y. Vernooy, M. Guo, and B.A. Hay, The Drosophila microRNA Mir-14 suppresses cell death and is required for normal fat metabolism. Curr Biol, 2003. **13**(9): 790–5.
162. Xu, P., M. Guo, and B.A. Hay, MicroRNAs and the regulation of cell death. Trends Genet, 2004. **20**(12): 617–24.
163. Yekta, S., I.H. Shih, and D.P. Bartel, MicroRNA-directed cleavage of HOXB8 mRNA. Science, 2004. **304**(5670): 594–6.
164. Yi, R., D. O'Carroll, H.A. Pasolli, Z. Zhang, F.S. Dietrich, A. Tarakhovsky, and E. Fuchs, Morphogenesis in skin is governed by discrete sets of differentially expressed microRNAs. Nat Genet, 2006. **38**(3): 356–62.
165. Ying, S.Y. and S.L. Lin, Intron-derived microRNAs–fine tuning of gene functions. Gene, 2004. **342**(1): 25–8.
166. Ying, S.Y. and S.L. Lin, Intronic microRNAs. Biochem Biophys Res Commun, 2005. **326**(3): 515–20.
167. Ying, S.Y. and S.L. Lin, Current perspectives in intronic micro RNAs (miRNAs). J Biomed Sci, 2006. **13**(1): 5–15.
168. Zeng, Y. and B.R. Cullen, Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. J Biol Chem, 2005. **280**(30): 27595–603.
169. Zeng, Y., R. Yi, and B.R. Cullen, Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. Embo J, 2005. **24**(1): 138–48.
170. Zhang, L., J. Huang, N. Yang, J. Greshock, M.S. Megraw, A. Giannakakis, S. Liang, T.L. Naylor, A. Barchetti, M.R. Ward, G. Yao, A. Medina, A. O'Brien-Jenkins, D. Katsaros, A. Hatzigeorgiou, P.A. Gimotty, B.L. Weber, and G. Coukos, microRNAs exhibit high frequency genomic alterations in human cancer. Proc Natl Acad Sci U S A, 2006. **103**(24): 9136–41.
171. Zhao, J.J., Y.J. Hua, D.G. Sun, X.X. Meng, H.S. Xiao, and X. Ma, Genome-wide microRNA profiling in human fetal nervous tissues by oligonucleotide microarray. Childs Nerv Syst, 2006. **22**(11): 1419–25.