Regulation of MicroRNA Expression and Abundance during Lymphopoiesis

Stefan Kuchen,1,10 Wolfgang Resch,1,10,* Arito Yamane,1,10 Nan Kuo,1 Zhiyu Li,1 Tirtha Chakraborty,2 Lai Wei,3 Arian Laurence,3 Tomoharu Yasuda,2 Siying Peng,2 Jane Hu-Li,4 Kristina Lu,5 Wendy Dubois,6 Yoshiaki Kitamura,7 Nicolas Charles,7 Hong-wei Sun,8 Stefan Muljo,4 Pamela L. Schwartzberg,5 William E. Paul,4 John O’Shea,3 Klaus Rajewsky,2 and Rafael Casellas1,9,*

1Genomics and Immunity, NIAMS, NIH, Bethesda, MD 20892, USA
2Immune Disease Institute and Department of Pathology, Harvard Medical School, Boston, MA 02115, USA
3Lymphocyte Cell Biology, NIAMS, NIH, Bethesda, MD 20892, USA
4Laboratory of Immunology, NIAID, NIH, Bethesda, MD 20892, USA
5Genetic Disease Research Branch, NHGRI, NIH, Bethesda, MD 20892, USA
6Laboratory of Genetics, NCI, NIH, Bethesda, MD 20892, USA
7Laboratory of Immune Cell Signaling, NIAMS, NIH, Bethesda, MD 20892, USA
8Biodata Mining and Discovery, NIAMS, NIH, Bethesda, MD 20892, USA
9Center of Cancer Research, NCI, NIH, Bethesda, MD 20892, USA
10These authors contributed equally to this work
*Correspondence: wresch@mail.nih.gov (W.R.), casellar@mail.nih.gov (R.C.)
DOI 10.1016/j.immuni.2010.05.009

SUMMARY

Although the cellular concentration of miRNAs is critical to their function, how miRNA expression and abundance are regulated during ontogeny is unclear. We applied miRNA-, mRNA-, and ChIP-Seq to characterize the microRNome during lymphopoiesis within the context of the transcriptome and epigenome. We show that lymphocyte-specific miRNAs are either tightly controlled by polycomb group-mediated H3K27me3 or maintained in a semi-activated epigenetic state prior to full expression. Because of miRNA biogenesis, the cellular concentration of mature miRNAs does not typically reflect transcriptional changes. However, we uncover a subset of miRNAs for which abundance is dictated by miRNA gene expression. We confirm that concentration of 5p and 3p miRNA strands depends largely on free energy properties of miRNA duplexes. Unexpectedly, we also find that miRNA strand accumulation can be developmentally regulated. Our data provide a comprehensive map of immunity’s microRNome and reveal the underlying epigenetic and transcriptional forces that shape miRNA homeostasis.

INTRODUCTION

MicroRNAs (miRNAs) are noncoding small RNAs that modulate the proteome of the cell by annealing to 3’ untranslated regions of cognate mRNAs and inhibiting protein translation and/or promoting mRNA instability (Bartel, 2004). Since their discovery in C. elegans (Lee et al., 1993; Wightman et al., 1993), miRNA orthologs and paralogs have been described in a variety of species, suggesting these regulatory RNAs are involved in basic cellular functions across the phyla (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001; Marson et al., 2008; Wightman et al., 1993). This view has been strengthened by the early embryonic lethality of mice deficient in miRNA processing factors (Bernstein et al., 2003; Chong et al., 2008; Kanellopoulou et al., 2005; Liu et al., 2004).

In the mammalian genome, miRNAs are encoded within introns of protein-coding genes or as independent entities transcribed either by RNA polymerase II (Rodriguez et al., 2004) or RNA polymerase III (Borchert et al., 2006). In some instances, groups of miRNAs are organized in genomic clusters processed from a single transcript. Because of their palindromic nature, miRNAs in nascent primary transcripts (pri-miRNAs) display a characteristic stem-loop structure that is recognized and cleaved in the nucleus by the Drosha-DGCR8 complex into 60–70 nucleotide (nt) precursor (pre) miRNAs. Once in the cytoplasm, pre-miRNAs are further processed by the RNase III endonuclease DICER into mature RNA fragments of ~22 nt in length, which are loaded into the RNA-induced silencing complex (RISC). Partial sequence complementarity between the 5’ end of the mature miRNA (6–8 nt seed region) and its target mRNA leads to downregulation of protein expression (Bartel, 2009; Doench and Sharp, 2004; Kim, 2005).

As is the case for nonhematopoietic tissues, lymphocytes and other cells of the immune system rely on miRNAs to effect lineage commitment, proliferation, migration, and differentiation (Taganov et al., 2007; Xiao and Rajewsky, 2009). In most cases, these activities are orchestrated by both ubiquitously expressed and hematopoietic-specific miRNA species (Basso et al., 2009; Landgraf et al., 2007; Merkerova et al., 2008; Monticelli et al., 2005; Neillson et al., 2007; Wu et al., 2007). Deletion or overexpression of these miRNAs impairs the immune system at various developmental stages (Chen et al., 2004; Li et al., 2007; O’Connell et al., 2008; Rodriguez et al., 2007; Thai et al., 2007; Ventura et al., 2008; Vigorito et al., 2007; Xiao et al., 2008). Similarly, conditional ablation of DICER or other miRNA processing factors...
results in a profound block of both B and T cell development (Cobb et al., 2005; Koralov et al., 2008; Muljo et al., 2005; O’Carroll et al., 2007). It is notable that these striking phenotypes are driven for the most part through small changes in the cellular concentration of key factors. In the B cell compartment for instance, miR-150 curtails the activity of the c-Myb transcription factor in a dose-dependent fashion over a narrow range of concentration of key factors. In the B cell compartment for instance, miR-150 curtails the activity of the c-Myb transcription factor in a dose-dependent fashion over a narrow range of miRNA and c-Myb concentrations (Xiao et al., 2007). Similarly, mass action seems to be the underlying principle behind miR-155 regulation of the B cell mutator Aid or miR-17 and miR-92-mediated inhibition of the tumor suppressor Pten and the proapoptotic Bim proteins (Dorsett et al., 2008; Teng et al., 2008; Ventura et al., 2008; Xiao et al., 2007; Xiao et al., 2008). These examples, which in hindsight explain the haploinsufficiency observed in Aid, cMyb, PTEN, and Bim heterozygous mice (Bouillet et al., 2001; Di Cristofano et al., 1998; Takizawa et al., 2008; Xiao et al., 2007), clearly demonstrate that the absolute cellular concentration of miRNAs is crucial at managing a cell’s proteome. Yet, how specific cell lineages establish miRNA concentrations upon differentiation remains to be defined.

Here, we use parallel sequencing to define chromatin modifications, the transcriptome, and miRNome of developing lymphocytes. This integrative approach reveals the epigenetic, transcriptional, and, indirectly, posttranscriptional mechanisms controlling miRNA cellular concentrations.

RESULTS

Deep Sequencing of Small RNAs
To determine miRNA abundance during lymphopoiesis, we microsequenced small RNAs (sRNAs) 18–30 nt in length from mouse hematopoietic progenitor cells (HPCs) and downstream T cell lineages including CD4^+CD8^+ thymocytes; splenic CD4^+ and CD8^+ T cells; ConA activated CD8^+ cells; ex-vivo differentiated Th (T helper type) 1, Th2, Th17, and iTreg (induced regulatory T) cells, as well as nTreg (natural regulatory T) cells; and Tfh (T follicular helper) cells, which were cell sorted from secondary lymphoid organs or spleens from immunized mice respectively (Table 1). In addition, essentially all stages of B cell ontogeny were examined, from progenitor bone marrow B cells (proB) to terminally differentiated plasma cells (Table 1). For comparative purposes, samples from other hematopoietic lineages (mast cells, basophils, neutrophils, dendritic cells, and NK cells), mouse embryonic stem cells (ESCs), and fibroblasts (MEFs), along with 11 adult tissues, were also included in the survey. Altogether, >300,000,000 sRNAs were sequenced. On average, four million sequence tags per library were aligned to the mouse genome with 100% identity and annotated as either miRNAs; noncoding (nc) RNAs (rRNAs, tRNAs, snoRNAs, or snRNAs); Piwi-interacting RNAs (piRNAs); small-coding RNAs (scRNAs); highly repetitive sequence; or unknown short RNAs (Figure 1A and Table S1 available online). With the exception of testes, which harbor a large number of piRNAs (Tam et al., 2008; Watanabe et al., 2008; Figure 1A), miRNAs were the most abundant RNA species identified, oscillating between 68% in germinal center (GC) B cells and 99.4% in the lung (Figure 1B). Interestingly, scRNAs, which are presumably mRNA breakdown products, were considerably enriched in GC B lymphocytes, and to a lesser extent, in other cells and tissues (Figure S1). Using mice expressing the E2-BclXL transgene, which blocks activation-induced cell death in B cells (Fang et al., 1996), we found that scRNAs are probably byproducts of mRNA degradation and thus may be analogous to the well-characterized internucleosomal DNA fragmentation occurring during apoptosis (Figure S1). We conclude that the large majority of sRNAs microsequenced belong to the miRNA family.

The Immune System’s MicroRNome
With the ultimate goal of defining the microRNome of developing lymphocytes, we next normalized the entire data set on the basis of the absolute number of miRNA reads sequenced in each sample. To validate this strategy, we quantified fold changes in miRNAs miR-191_5p, miR-21_5p, miR-25_3p, miR-128-1_3p, miR-128-2_3p, let7c-1_5p, let7c-2_5p, and let-7e_5p between brain and activated B cell samples by using LNA qPCR. Reassuringly, we observed a high degree of correlation between LNA and miRNA-seq (R^2 = 0.995, Figure S2A). Using TaqMan technology, we also analyzed eight miRNAs in various B and T cell populations as well as HPCs and found again an overall agreement between TaqMan and deep sequencing (p = 0.76; CI.95 = [0.66, 0.83]; p < 0.0001, Figure S2B). Finally, biological and technical replicates from mature resting B cells and neutrophils respectively showed significant reproducibility between miRNA-seq runs (p > 0.94; p < 0.0001 for all replicates, Figure S2C). These analyses demonstrate that deep sequencing can reliably profile miRNA abundance relative to the total pool of miRNAs; thus our normalization protocol permits direct comparison of miRNA expression between different cells and tissues.

Of the 600 mature mouse miRNAs (miRBase 13), 353 were detected at more than 100 sequence tags per million (TPM) in at least one of the 40 tissues and cell types analyzed (Table S2). Moreover, using a modified version of the miRDeep algorithm (Friedländer et al., 2008, see Experimental Procedures), we isolated 18 putative miRNAs, the majority of which showed phylogenetic conservation (Data S1). Each individual tissue or cell type expressed a mean of 102 ± 28 miRNAs at greater than 100 TPM and 162 ± 43 when the cutoff was set at 25 TPM. In agreement with previous studies (Landgraf et al., 2007), hierarchical clustering of miRNA profiling both paralleled known developmental histories and clearly separated cells of the immune system from other tissues (Figure 2A). Unexpectedly, expression of the let-7 family of miRNAs also partitioned hematopoietic from nonhematopoietic cells. For instance, let-7f was predominant in the immune system, comprising in some cases up to 60% of all let-7 sequences, whereas its expression in nonhematopoietic cells was statistically lower (p < 0.0001; Wilcoxon rank-sum test, Figure 2B). Conversely, let-7c abundance was substantially greater in nonhematopoietic cells (Figure 2B). Other let-7 family members also showed differential expression in these two compartments (data not shown). Thus, profiling and hierarchical clustering of miRNAs set apart hematopoietic lineages from other cells and tissues.

ESCs displayed a large number of specific miRNAs (151), followed by brain (52) and testes (41) (Figure 2A; Table S2 and Data S2). In the immune system, neutrophils, basophils, as well as T and B cells (considered as a single group) also displayed
a substantial number of preferentially expressed miRNAs (Figure 2A). In addition to miR-223, which has been previously ascribed to the myeloid lineage (Chen et al., 2004), neutrophils expressed high amounts of miR-26 (a and b), miR-744, miR-340, and eight other miRNAs (Figure 2C). In striking contrast to ESCs, hematopoietic progenitor cells displayed few unique miRNAs, illustrated in Figure 2D by miR-193b. From basophils and mast cells we characterized a highly specific miRNA (1073496_chr3) located in intron 8 of the CPA3 gene, which encodes the basophil and mast cell-secreted protease carboxypeptidase A (Figure 2D). We also uncovered four miRNAs with preferential expression in NK cells. Notably, these miRNAs displayed nearly identical expression profiles, which, in addition to NK cells, included lower expression in Tfh cells, ConA-activated CD8+ T cells, Th1 cells, and LPS+IL-4 stimulated B lymphocytes (Figure 2D and Data S1). This feature implies that a common regulatory pathway might drive expression of these miRNAs.

A total of 49 miRNAs were preferentially upregulated in lymphocytes (Figure 3A and Table S3). As formerly shown (Chen et al., 2004), the miR-181 family was highly abundant in developing bone marrow B cells and thymocytes (Figure 3A). In CD4+CD8+ T cells, miR-181 miRNAs comprised 13.7% (137,024 TPM) of the microRNome, which is in good accord with published estimates (15.6%; Neilson et al., 2007). miR-128, previously implicated in neural specification and malignancy, was even more dominant in CD4+CD8+ T cells (54,134 TPM compared to 27,568 TPM in the brain, Figure 3B). In B cells, the most abundant miRNAs were miR-320 and miR-191, which comprised as much as 10% (109,145 TPM) of all miRNAs in Table 1. Tissues and Cell Types

<table>
<thead>
<tr>
<th>Cell Type</th>
<th>Species</th>
<th>Organ</th>
<th>Sorting, Purification, and Culture Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>ProB cells</td>
<td>Mouse</td>
<td>Bone marrow</td>
<td>B220+IgM CD43+CD25+</td>
</tr>
<tr>
<td>PreB cells</td>
<td>Mouse</td>
<td>Bone marrow</td>
<td>B220+IgM CD43+CD25+</td>
</tr>
<tr>
<td>Immature B cells</td>
<td>Mouse</td>
<td>Spleen</td>
<td>B220+IgM-CD21/CD35b</td>
</tr>
<tr>
<td>Mature B cells</td>
<td>Mouse</td>
<td>Spleen</td>
<td>B220+IgM-CD21/CD35b</td>
</tr>
<tr>
<td>Marginal zone B cells</td>
<td>Mouse</td>
<td>Spleen</td>
<td>B220+IgM-CD21/CD35b</td>
</tr>
<tr>
<td>B1 B cells</td>
<td>Mouse</td>
<td>Peritoneal cavity</td>
<td>B220+IgM-6h</td>
</tr>
<tr>
<td>Germlinal center B cells</td>
<td>Mouse</td>
<td>Lymph nodes (immunized)</td>
<td>CD19-CD95+</td>
</tr>
<tr>
<td>Plasma cells</td>
<td>Mouse</td>
<td>Lymph nodes</td>
<td>B220-CD138h from IL6 transgenic mice</td>
</tr>
<tr>
<td>LPS+IL4 B cells</td>
<td>Mouse</td>
<td>Spleen</td>
<td>CD43 depleted; 50 μg/ml LPS, 2.5ng/ml IL4, 72 hs</td>
</tr>
<tr>
<td>LPS+α-β-dextran B cells</td>
<td>Mouse</td>
<td>Spleen</td>
<td>CD43 depleted; 50 μg/ml LPS, 2.5ng/ml α-β-dextran, 72 hs</td>
</tr>
<tr>
<td>Mast cells</td>
<td>Mouse</td>
<td>Bone marrow</td>
<td>20 ng/ml IL3, 20ng/ml SCF, 8 weeks</td>
</tr>
<tr>
<td>Basophils</td>
<td>Mouse</td>
<td>Bone marrow</td>
<td>20 ng/ml IL3, sorted at day 10, CD49b-FccR1+CD11b+Kit1-</td>
</tr>
<tr>
<td>Neutrophils</td>
<td>Mouse</td>
<td>Peritoneal cavity (casein)</td>
<td>Percoll gradient, 88% Ly-6G+ purity</td>
</tr>
<tr>
<td>Dendritic cells</td>
<td>Mouse</td>
<td>Bone marrow</td>
<td>20 ng/ml GM-CSF, non-adherent, 12 days</td>
</tr>
<tr>
<td>Macrophages</td>
<td>Mouse</td>
<td>Bone marrow</td>
<td>10 ng/ml GM-CSF, 10ng/ml IL3, nonadherent, 14 days</td>
</tr>
<tr>
<td>NK cells</td>
<td>Mouse</td>
<td>Spleen (RA31-)</td>
<td>1000 U/ml IL2, adherent, 8 days</td>
</tr>
<tr>
<td>Hematopoietic progenitor cells</td>
<td>Mouse</td>
<td>Bone marrow (S-FU)</td>
<td>20 ng/ml IL3, 50 ng/ml IL6, 50 ng/ml SCF, 7 days, 60% c-Kit+Scal+Lin- purity</td>
</tr>
<tr>
<td>Double-positive T cells</td>
<td>Mouse</td>
<td>Thymus</td>
<td>CD3+CD4+CD8+</td>
</tr>
<tr>
<td>Naive CD4+ T cells</td>
<td>Mouse</td>
<td>Lymph node/Spleen</td>
<td>CD4+CD62L+CD44iwiCD25+</td>
</tr>
<tr>
<td>Naive CD8+ T cells</td>
<td>Mouse</td>
<td>Spleen</td>
<td>CD3+CD8+CD4-</td>
</tr>
<tr>
<td>ConA CD8+ T cells</td>
<td>Mouse</td>
<td>Spleen</td>
<td>5 μg/ml Concanavalin A, 5 days</td>
</tr>
<tr>
<td>Th1 T cells</td>
<td>Mouse</td>
<td>Spleen (sorted CD4+)</td>
<td>Protocol from Wei et al. (2009)</td>
</tr>
<tr>
<td>Th2 T cells</td>
<td>Mouse</td>
<td>Spleen (sorted CD4+)</td>
<td>Protocol from Wei et al. (2009)</td>
</tr>
<tr>
<td>Th17 T cells</td>
<td>Mouse</td>
<td>Spleen (sorted CD4+)</td>
<td>Protocol from Shi et al. (2008)</td>
</tr>
<tr>
<td>T follicular helper T cells</td>
<td>Mouse</td>
<td>Spleen (immunized)</td>
<td>CD4+CD44hiCXCR5hi</td>
</tr>
<tr>
<td>Natural regulatory T cells</td>
<td>Mouse</td>
<td>Lymph node/spleen</td>
<td>CD4+CD25+</td>
</tr>
<tr>
<td>Induced regulatory T cells</td>
<td>Mouse</td>
<td>Spleen (sorted CD4+)</td>
<td>Protocol from Wei et al. (2009)</td>
</tr>
<tr>
<td>ESCs</td>
<td>Mouse</td>
<td>Blastocysts</td>
<td>Described in Marson et al. (2008)</td>
</tr>
<tr>
<td>Mouse embryonic fibroblasts</td>
<td>Mouse</td>
<td>E13.5 (C57B6)</td>
<td>Passage 2</td>
</tr>
<tr>
<td>Naive B cells</td>
<td>Human</td>
<td>Tonsil</td>
<td>CD19+CD38hiCD27+</td>
</tr>
<tr>
<td>Memory B cells</td>
<td>Human</td>
<td>Tonsil</td>
<td>CD19+CD38hiCD27+</td>
</tr>
<tr>
<td>Pregerminal Center B cells</td>
<td>Human</td>
<td>Tonsil</td>
<td>CD19+CD38+CD77-IgD</td>
</tr>
<tr>
<td>Centroblasts</td>
<td>Human</td>
<td>Tonsil</td>
<td>CD19+CD38+CD77-IgD</td>
</tr>
<tr>
<td>Centrocytes</td>
<td>Human</td>
<td>Tonsil</td>
<td>CD19+CD38+CD77-IgD</td>
</tr>
<tr>
<td>Plasma cells</td>
<td>Human</td>
<td>Tonsil</td>
<td>CD19+CD38hiCD27hiIgD</td>
</tr>
</tbody>
</table>

Please cite this article in press as: Kuchen et al., Regulation of MicroRNA Expression and Abundance during Lymphopoiesis, Immunity (2010), doi:10.1016/j.immuni.2010.05.009
plasma cells and up to 5% in bone marrow and some peripheral B cells, respectively (Figure 3C). By comparison, the highest expression of miR-155 in activated B cells was 3,417 TPM (Table S2). Other notable examples were miR-147, miR-31, and miR-15b, which showed predominant expression in Th1, GC B lymphocytes respectively (Figure 3C). We also isolated several uncharacterized miRNAs in lymphocytes, including 482234_chr15, which was confined to IgMhi immature B cells, and 145561_chr11, present in LPS+IL-4-activated B cells, iTreg cells, Th17 cells, and ESCs (Figures 3B and 3C). Both miRNAs were phylogenetically conserved and intrinsic to noncoding mRNAs (Figure 3, schematics). On the basis of these results, we conclude that miRNA-Seq reliably identifies the microRNome signature of hematopoietic cells, including lymphocytes.

Epigenetic Regulation of miRNA Expression in Lymphocytes

To examine how miRNA expression is regulated during lymphopoiesis, we first mapped histone modifications at high resolution across the entire genome of HPCs, proB, preB, mature resting, and LPS+IL-4 activated B cells. We concentrated on histone H3, lysine 4 trimethylation (H3K4me3), which demarcates both gene promoters and transcriptional activity; H3K27me3, which is associated with transcriptional inhibition; and H3K36me3, which follows transcriptional elongation. These data were complemented with similar epigenetic analyses of ESCs (Marson et al., 2008), as well as naive CD4+ T cells and downstream-differentiated populations Th1, Th2, Th17, iTreg, and nTreg cells (Wei et al., 2009). For a more comprehensive view of the epigenetic landscape at and around miRNA genes, we mapped 32 additional histone modifications in mature resting and LPS+IL-4 activated B cells (H2AK9Ac, H2BK5Ac, H2BK5me1, H2BK12Ac, H2BK20Ac, H2BK120Ac, H3K4me1, H3K4me2, H3K9Ac, H3K9me1, H3K9me2, H3K9me3, H3K4Ac, H3K4me1, H3K4me2, H3K9Ac, H3K9me1, H3K9me2, H3K9me3, H3K4Ac, H3K14Ac, H3K18Ac, H3K23Ac, H3K27Ac, H3K36Ac, H3K27me1, H3K27me2, H3K36me1, H3K36me2, H3K36me3, H3K79me1, H3K79me2, H3K79me3, H4K5Ac, H4K8Ac, H4K12Ac, H4K16Ac, H4K91Ac, H4K20me1, and H3K20me3), together with the histone variant H2A.Z, RNA polymerase II (PolII), and the histone variant H2B.K.120Ac, H3K4me1, H3K4me2, H3K9Ac, H3K9me1, H3K9me2, H3K9me3, H3K4Ac, H3K14Ac, H3K18Ac, H3K23Ac, H3K27Ac, H3K36Ac, H3K27me1, H3K27me2, H3K36me1, H3K36me2, H3K36me3, H3K79me1, H3K79me2, H3K79me3, H4K5Ac, H4K8Ac, H4K12Ac, H4K16Ac, H4K91Ac, H4K20me1, and H3K20me3), together with the histone variant H2A.Z, RNA polymerase II (PolII), and the histone acetyltransferase CBP-p300 (Table S4). The raw epigenetic data of miRNA (±5 kb) is provided as Data S3. This bedgraph file can be loaded and viewed using the UCSC browser (http://genome.ucsc.edu/).

On the basis of their expression profiles during B and T cell development, we broadly classified miRNAs as (1) inactive, (2) active, (3) poised, or (4) induced (Figure 4). As expected, miRNA genes not expressed during lymphopoiesis (e.g., the heart-specific miR-383) displayed repressive modifications such as H3K27me3, H3K9me3, H3K27me2, and/or H4K20me3 (Wang et al., 2008; Figure 4A and Table S4). In keeping with their lack of transcription, we detected little or no PolII at promoter regions of these genes (Figure 4A and Table S4). Alternatively, miRNA genes fully active across development (e.g., let-7g) displayed chromatin modifications previously linked to actively transcribed genes ([Wang et al., 2008], Figure 4B and Table S4). We point out that whereas lack of posttranscriptional mechanisms prevent accumulation of let-7 mature miRNAs in ESCs [Viswanathan et al., 2008 and Figure 4B, bar graph]; the let-7 genes are nonetheless fully expressed therein.)

Among miRNAs confined to specific stages of B and/or T cell development, our epigenetic analysis revealed at least two distinct subsets. As exemplified by miR-155 and 145561_chr11, one group was H3K27 demethylated in HPCs or resting cells, yet showed little or no expression at these early stages of...
development (Figure 4C, Table S4, and Data S3). Cell activation, however, brought about a dramatic increase in miRNA expression along with a high degree of activating modifications as well as PolII and p300 recruitment. These miRNA genes thus appear to be poised for activation early during ontogeny. On the other hand, other hematopoietic-specific miRNAs, illustrated in Figure 2. miRNA Signatures of the Mouse Immune System, Embryonic Cells, and Adult Tissues

(A) Heat map showing expression of 375 miRNAs (detected at >100 TPM) in mouse hematopoietic and nonhematopoietic cells based on hierarchical clustering analysis. miRNAs enriched in a particular sample are depicted with red bars, whereas depleted miRNAs are depicted with green bars. B and T cell populations outlined in Table 1 were combined into a single group.

(B) Percentage of let-7a (gray line), let-7c (red line), and let-7f (blue line). The total let-7 sequence tags were set to 100%. Let-7f and let-7c miRNA abundance is inversely proportional in hematopoietic cells (light-blue box), whereas roughly equivalent in nonhematopoietic cells and tissues (light-red box).

(C) Expression profiles of miRNAs enriched in neutrophils. Two independently processed neutrophil samples (1 and 2) are shown and the highest TPM value of the two is given in parenthesis.

(D) Examples of miRNAs predominantly expressed in HPCs, mast cells, and NK cells. For uncharacterized miRNAs, their genomic location is schematized. Orientation of miRNAs and genes is represented with red and black arrowheads, respectively. Mammalian phylogenetic conservation is depicted with dense black bar graphs based on UCSC PhastCons30Way algorithm.
by miR-139 and miR-147, were more tightly regulated in that they retained H3K27me3 up until full miRNA gene expression was invoked (Figures 4D, Table S4, and Data S3). Transcription was then accompanied by an increase in active chromatin marks such as H3K4me3 and H3K36me3. Importantly, in the case of miR-139, which was expressed in proB and preB cells, transcriptional downregulation in mature resting and activated B cells was accompanied by a return of H3K27me3 (Figure 4D), strengthening the notion that targeted expression of this miRNA subset is tightly regulated. On the basis of these results, we conclude that lymphocyte-specific miRNAs differ as to when the repressive mark deposited by polycomb group proteins, H3K27me3, is selectively lost. The potential functional impact of such distinction in the context of B and T cell differentiation is discussed below.

Transcriptional Regulation of miRNA Expression in Lymphocytes

One corollary of our epigenetic analysis is that, at least for some miRNAs, changes in miRNA cellular concentration during lymphopoiesis are probably determined by gene transcription. This view is consistent with the observation that clustered miRNAs in the genome display fairly similar expression profiles (Figure 5A; Baskerville and Bartel, 2005; Ruby et al., 2007). The rationale being that proximally located miRNAs are encoded by polycistronic group proteins, H3K27me3, is selectively lost. The potential functional impact of such distinction in the context of B and T cell differentiation is discussed below.

Ectopically Expressed mRNA Targets Influence miRNA Abundance

The above results uncovered examples where one of two miRNA strands more closely parallels changes in miRNA gene expression (for instance, Figure 5D), implying that mechanisms other than transcription and/or posttranscriptional processing may influence final miRNA cellular abundance. To systematically address this issue, we quantified expression of spliced primary transcripts by miRNA-seq and compared mRNA RPKM to miRNA TPM values in six stages of B cell development (HPCs, proB, preB, resting mature, LPS+IL-4 activated, and GC B cells). On the basis of a Spearman’s rank correlation coefficient > 0.7, we found coordinate expression of spliced primary transcripts and mature miRNAs in 15 cases out of 54 (27%, Figure 5B and Table S5). Within this correlated group, we found examples of miRNAs intronic to coding genes illustrated by miR-139 and miR-152 (Figure 5C), or embedded within noncoding RNAs, as in the case of miR-107 and miR-362 (Table S5). We conclude that during B cell development, fluctuations in expression of one-quarter of miRNAs are correlated to spliced primary transcript amounts.
than 3-fold strand bias (Figure 6A, enclosed with dotted lines), and the remaining 47% fall somewhere in between these two extremes. To determine whether this strand distribution was evolutionarily conserved, we deep-sequenced the human B cell microRNome from six tonsillar populations (Data S4) and compared the 5p:3p ratio between mouse and human for miRNAs with exact sequence conservation (highlighted in red in Figure 6A). We found extensive 5p:3p correlation between the two species ($r = 0.94; CI.95 = [0.89, 0.96]; p < 0.0001$, Figure 6B), indicating miRNA strand abundance might be under similar mechanistic constraints across species.

The uneven accumulation of 5p and 3p miRNA strands has been explained by the functional asymmetry principle, which postulates that during processing of miRNA duplexes, the strand preferentially loaded into the RISC complex is the one whose 5’ end is less tightly paired to its complement (Schwarz et al., 2003). The ejected passenger strand or miRNA* is degraded, and thus in general less abundant. Supporting the functional asymmetry principle, a systematic comparison of 5p-3p strand concentrations against the calculated stability of the 5’ and 3’ ends of the miRNA duplex showed an overall significant correlation ($r = 0.53; CI.95 = [0.42, 0.61]; p < 0.0001$, Figure 6C). This finding, which is consistent with previous studies (Schwarz et al., 2003; Tomari et al., 2004), argues that cellular 5p-3p distributions can be partially explained by the propensity of the miRNA duplex ends to fray. Yet, we found multiple instances where miRNA strands showed differential expression profiles across the various tissues and cell types examined. A case in point is miR-30c2, whose 3p strand is undetected in hematopoietic cells, but expressed at levels comparable to its complement, miR-30c2_5p, in somatic cells and tissues (Figure 6D, left graph). Another example is the miR-150_5p _3p pair, which is expressed in resting immature and mature B cells as 3p > 5p, whereas this pattern is reversed in the T cell compartment (Figure 6D, right graph). Other examples are miR-139, miR-329, and miR-30e (Table S2).

At least one way the above observations can be explained is through cognate mRNA target abundance, which is expected
to vary throughout ontogeny and could in principle impact the stability and/or half-life of miRNA-RISC complexes. To explore this possibility, we expressed in A70 proB cells miR-30_5p complementary target sequences (mirTs; Gentner et al., 2009), which are recognized by the seed domain of the entire miR-30 family of miRNAs. We found that mirT expression affected in highly specific fashion the steady-state abundance of miR-30d_5p and miR-30e_5p, the two miR-30 members expressed at significant amounts in A70 cells (Figure S4A). Specifically, miR-30_5p amounts declined relative to control upon cognate mRNA expression: 463 versus 3320 TPM for miR-30d_5p and 330 versus 1237 TPM for miR-30e_5p (Figure S4A). Attesting to the high specificity of the mirT-miRNA interaction, no other endogenous miRNAs (including miR-30_3p arms) were drastically affected by miR-30_5pT expression. Analogously, expression of miR-30_5pTs in 293T cells resulted in specific depletion of 5p arms from the two expressed miR-30 members: miR-30e_5p (2001 versus 7693. TPM) and miR-30a_5p (2840 versus 868 TPM, Figure S4B), indicating that downregulation of miR-30_5p by mRNA targets is conserved between mice and humans. We confirmed these results by targeting 4 additional endogenous miRNAs in 293T cells. Interestingly, we found examples of both specific upregulation (miR-744_5p) and downregulation (miR-25_3p, miR-128_3p, miR-191_5p) of miRNAs upon cognate mRNA expression (Figure S4C). We conclude that the steady-state abundance of miRNA strands can be influenced by ectopic expression of complementary mRNA sequences. As proposed below, this feature might help explain the changes in 3p:5p strand ratios observed during development.

**DISCUSSION**

The hematopoietic microRNome can be set apart from that of somatic cells and tissues by a distinctive miRNA signature, differential expression patterns of let7, and a unique 5p:3p
distribution for selected miRNAs. Previous studies uncovered miRNAs circumscribed to or enriched in the immune system, such as miR-150, miR-155, miR-223, miR-146, and the miR-181 family (Barski et al., 2009; Basso et al., 2009; Chen et al., 2004; Landgraf et al., 2007; Merkerova et al., 2008; Monticelli et al., 2005; Neilson et al., 2007; Taganov et al., 2006; Wu et al., 2007; Zhang et al., 2009). Our comparative cluster analysis corroborated these findings and further assigned previously characterized miRNAs to the hematopoietic system: miR-15b (nTreg cells), miR-26a/b (neutrophils), miR-28 (GC B cells), miR-107 (basophils), miR-320 and miR-148a (plasma cells), miR-128 (CD4+CD8+ cells), miR-128-2 (CD4+CD8+ cells), miR-221 and miR-222 (Tfh-basophils), miR-147 (Th1 cells), miR-182 (NK cells), and miR-193b (HPCs). Of note, both miR-28 and miR-148a were also highly specific for human GC and plasma cells, respectively (Figure S6). In addition, we identified 14 mouse miRNAs that show preferential expression in various cells of the immune system 

In the context of gene discovery, it is important to point out that of the >300,000,000 sequence reads analyzed, as many as 2,406 putative mouse miRNAs were predicted on the basis of a probabilistic model of miRNA biogenesis (Friedländer et al., 2008; data not shown). Of these, however, only a small fraction (18 in total) was detected at amounts greater than 100 TPM. It is interesting to note therefore that the overwhelming majority of putative miRNAs uncovered by our analysis appear to represent low-abundant hairpin species that might on occasion be processed by the RNAi pathway. Admittedly, we cannot rule out that some of these putative miRNAs are expressed at considerable levels in tissues not examined in our survey. Furthermore, the arbitrary 100 TPM threshold is not a predictor of physiological activity, which in principle could still be significant at very low levels of miRNA expression. In light of our findings however, it is reasonable to propose that most mouse and human miRNAs have already been characterized. This idea is supported by the fact that recent large-scale cloning efforts yielded few uncharacterized, mostly low-abundant, miRNAs (Basso et al., 2009; Landgraf et al., 2007).

The combining high-throughput miRNA-seq, mRNA-seq, and ChIP-seq provides an unprecedented view of the various regulatory steps that shape miRNA expression and abundance during
lymphopoiesis. At the epigenetic level, the methylation status of H3K27 partitioned lymphocyte-specific miRNA genes into two distinct subsets. Those belonging to the miR-139 and miR-147 group were found to retain the repressive H3K27me3 mark throughout development up until miRNA gene transcription was elicited. miRNA genes of the miR-155/145561_chr11 group were H3K27 demethylated at earlier stages of development, and thus they appeared to be poised for activation prior to full transcription. The rationale for this subdivision might be provided by the spatiotemporal context in which these miRNAs are expressed. miR-147 for instance is induced as naive CD4+ T cells differentiate into the Th1 cell lineage. This process of fate determination relies on positive epigenetic reprogramming and expression of key factors, and negative regulation of competing pathways driving differentiation to other T helper cell types (Wei et al., 2009). Similarly, miR-139 derepression occurs as part of another fate determination step, from pluripotent hematopoietic progenitor cells to pro-B cells. Although the precise role of miR-139 and miR-147 remains to be determined, the expectation would be that miRNAs involved in symmetric or progressive lymphocyte lineage commitment would be tightly regulated by robust mechanisms. One such mechanism may be polycomb group-mediated H3K27 trimethylation, which might ensure induction of cell-fate determination only at the appropriate time.

Our data have also revealed lymphocyte-induced miRNA genes that appear to be epigenetically poised for transcription early in development. These miRNA genes are associated with some activating histone marks, are H3K27me3 depleted, recruit low levels of PolII, but lack H3K36me3 or H3K79me2—hallmarks of polymerase elongation. This subinduced chromatin state is reflected by the low, but detectable, miRNA expression prior to full gene transcription. For miR-155, 145561_chr11, and other miRNAs under this category, full expression occurs as resting lymphocytes are exposed to mitogens and cytokines typically released during the immune response. The benefits of maintaining critical miRNA genes like miR-155 in a preexisting subinduced state might be best understood in the cadre of viral or bacterial infections, which require rapid T and B cell responses. On the basis of these considerations, it will be important to determine whether miRNAs strictly dependent on H3K27me3 for expression are involved in lymphocyte developmental decisions, while H3K27me3-independent ones drive B and T cell activation. Ongoing studies indicate that this might be indeed the case.

Transcriptionally, we have shown that expression of at least one-quarter of mature miRNAs closely follows spliced primary transcripts as determined by mRNA-Seq. Because mRNAs are stabilized by polyadenylation and other mechanisms, this is likely to be an underestimate, e.g., the correlation between mature miRNAs and unspliced pri-miRNAs is expected to be more significant. Regardless of the precise number, the results were nonetheless unexpected in light of the fact that studies with cell lines fail to find any such correlation (Thomson et al., 2006). Transformed cells, however, display a global reduction in mature miRNAs (Lee et al., 2008; Lu et al., 2005), presumably from a general deficiency in miRNA processing. Alternatively, in primary cells several posttranscriptional mechanisms are bound to promote deviations between transcription rate and mature miRNA expression, such as efficiency of miRNA processing, miRNA editing, or miRNA nuclear transport. In addition to these, we have found that ectopically expressed mRNA targets can both increase or decrease a particular miRNA pool. Whether endogenous mRNAs can likewise influence fluctuations in miRNA abundance remains to be determined. Nevertheless, it is reasonable to propose that an increase in the absolute number of miRNA targets may impact the balance and/or half-life of “free” miRNA, mRNA-RISC, and miRNA-RISC-mRNA complexes. We find this scenario intriguing because it raises the possibility that under physiological conditions, miRNAs and mRNAs regulate each other’s homeostasis.

In summary, our studies reveal some of the epigenetic, transcriptional, and posttranscriptional strategies that help orchestrate cellular abundance of miRNAs during lymphopoiesis. The hematopoietic miRNA signatures provided by the data represent a valuable resource that will help guide future gene-targeting experiments of individual miRNAs. In this respect, a major challenge in the field has been the identification of critical miRNAs: miRNA targets driving developmental decisions. A strategy to solve this problem is based on the observation that cellular concentration of miRNA targets fluctuates as a function of cognate miRNA expression (e.g., miR-150c-Myb [Xiao et al., 2007]). In principle, by applying microsequencing and bioinformatics to a large number of developmental stages, it should be possible to predict functional miRNA:mRNA target pairs. Our preliminary studies using the miRNA- and mRNA-Seq data from B cell development support this view. We anticipate that genomic approaches such as this will help unravel how miRNAs regulate development and effector functions of the immune system.

**EXPERIMENTAL PROCEDURES**

**RNA Isolation and Quality Control**

Sorted or cultured cells were collected by centrifugation, dissolved in TriZol at a concentration of 5 to 20 × 10⁶ cells per milliliter, and stored at −80°C. Total RNA was isolated in accordance with the manufacturer’s recommendations. The RNA integrity number (RIN) for all samples was assessed from 100 ng of total RNA with Agilent RNA 6000 Nano Kit and Bioanalyzer 2100 (Agilent Technologies). The RIN of all mouse and human cell populations was at least 9, and the RIN of tissues was at least 8 except for pancreas (RIN 7.3) and skin (RIN 7).

**Small RNA Profiling by Illumina Deep-Sequencing**

Small RNA sample preparation was done in accordance with Illumina’s protocol. In brief, 5’ and 3’ adapters were sequentially ligated to small RNA of 18–30 bases gel purified from 5–10 μg total RNA. Adapter-ligated small RNA was reverse transcribed, amplified by 15 PCR cycles with high-fidelity Phusion polymerase (Finnzymes), and sequenced on a Genome Analyzer (Illumina) in accordance with the manufacturer’s instructions.

**Epigenetic Analysis**

Cells were fixed with 1% paraformaldehyde at 37°C for 10 min and either MNase-treated for ChIP or sonicated for Polymerase II and p300 IP. Precipitated DNA fragments were processed in accordance with Illumina’s protocol and sequenced on a Genome Analyzer with manufacturer’s instructions. During analysis, short sequence reads were trimmed to 25 nts and aligned to the mouse genome with either ELAND or Bowtie. Comparison data for mESC and T cell populations were obtained from public databases (Marson et al., 2008; Wei et al., 2009). Uniquely aligned reads were analyzed by SICER (Zang et al., 2009) with an expectation value E of 50 in a random background model. Reads on significant islands as defined by SICER were normalized to the total number of reads on islands. Downstream analysis was carried out in R.
Quantitative RT-PCR
RNA samples used for quantitative RT-PCR were isolated with the mirVana miRNA isolation kit (ABI, part No: 1561) in accordance with the manufacturer’s instructions. cDNA generated from 10 ng of total RNA by the TaqMan microRNA reverse transcription kit (ABI, part No: 4366596) was applied for each TaqMan qRT-PCR reaction with TaqMan 2 × Universal PCR Master Mix (ABI, part No: 4324018). All TaqMan qRT-PCR reactions were performed and analyzed by StepOnePlus Real-time PCR system (ABI) with the following cycle conditions: 95 °C for 10 min, 1 cycle; 95 °C for 15 s and 60 °C for 60 s, 50 cycles. The expression of individual miRNA was normalized and expressed as a percentage relative to U6 with the following formula: fold induction = 2^{ΔΔCt}, where ΔΔCt = C(targen) − C(U6).

For a comparison of brain samples to activated B cells, total RNA was sent to Exicon for commercial miRCURY™-based microRNA quantification of selected microRNAs (mmu-let-7c, mmu-let-7e, mmu-miR-21, mmu-miR-25, mmu-miR-128, and mmu-miR-191).

ACCESSION NUMBERS
All sequence data are available in the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE21630.

SUPPLEMENTAL INFORMATION
Supplemental Information includes three figures, five tables, four data collections, and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.immuni.2010.05.009.

ACKNOWLEDGMENTS
We thank members of the Casellas lab for discussions; J. Newman and R. Young for mouse ESCs; J. Simone for cell sorting; G. Gutierrez for technical assistance with the genome analyzer; B. Wold for the miRNA-seq protocol; and L. Naldini for the LV-SFFV lentiviral vector. This work was supported in part by the Intramural Research Program of NIAMS-NIH. S. K. was supported by the Swiss Foundation for Grants in Biology and Medicine.

Received: October 18, 2009
Revised: March 22, 2010
Accepted: April 8, 2010
Published online: June 3, 2010

REFERENCES


