Distinct microRNA signatures in human lymphocyte subsets
and enforcement of CD4+ T cell naïve state by miR-125b

Riccardo L. Rossi¹, Grazisa Rossetti¹, Lynn Wenandy¹, Serena Curti¹, Anna Ripamonti¹, Raoul J. P. Bonnal¹, Roberto Sciarretta Birolı¹, Monica Moro¹, Maria C. Crosti¹, Paola Gruarin¹, Stefano Maglie¹, Francesco Marabita¹, Debora Mascheroni¹, Valeria Parente¹, Mario Comelli², Emilio Trabucchi³, Raffaele De Francesco¹, Jens Geginat¹, Sergio Abrignani¹, Pensilno Masa⁰ & Massimiliano Pagani¹

1. Istituto Nazionale Genetica Molecolare (INGM), Via F. Sforza 28, Milan 20122, Italy
2. Dipartimento di Scienze Sanitarie Applicate, Università di Pavia, Via A. Bassi, 21, Pavia 27100, Italy
3. Pio Albergo Trivulzio, Via Trivulzio 15, Milan 20146, Italy
4. IRCCS Ca’Granda Ospedale Maggiore Policlinico, Via F. Sforza 28, Milano 20122, Italy
5. These authors contributed equally to this work

Correspondence should be addressed to: abrignani@ingm.org; pagani@ingm.org
MicroRNAs are small non-coding RNAs that regulate gene expression post-transcriptionally. Here we applied microRNA profiling in seventeen human lymphocyte subsets to identify microRNA signatures that were distinct among various subsets and different from those of mouse lymphocytes. miR-125b, one of the signature microRNAs of naïve CD4+ T cells, regulated expression of genes involved in T-cell differentiation i.e., IFNG, IL2RB, IL10RA and PRDM1. Expression of synthetic miR-125b RNA and lentiviral vectors encoding pre-miR-125b in naïve lymphocytes inhibited differentiation to effector cells. Our data provide an atlas of microRNA expression in human lymphocytes, define subset-specific signatures and their target genes, and indicate T-cell naïve state is enforced by microRNAs.
INTRODUCTION

MicroRNAs (miRNA) are highly conserved non-coding single-stranded small RNA molecules that control gene expression post-transcriptionally by binding the 3’UTR of target mRNAs. This results in degradation or translation inhibition of target mRNAs. miRNAs are expressed in a tissue-specific manner, and changes in their expression within a tissue have been associated with several human diseases, including cancer, inflammation, and chronic viral infections.

Although differentiation and functions of T lymphocytes have been thoroughly investigated through comprehensive gene expression analysis and several transcription factors have been identified as essential for the differentiation of naïve CD4⁺ T cells into different effector subsets, the molecular basis of CD4⁺ T helper (T_H) cell differentiation and of the process through which T cells acquire their effector function is incompletely understood.

miRNAs are differentially expressed during immune cells development and function, and their importance has been highlighted by experiments showing that disruption of Ago2 (the core of RNA-Induced Silencing Complex) in hematopoietic cells results in impaired B cell differentiation and that conditional deletion of Dicer in thymocytes induces a strong reduction of T lymphocytes in the thymus and periphery. Moreover, several studies suggest a role for specific miRNAs in lymphocyte biology. miR-181 modulates T cell antigen receptor sensitivity, miR-150 regulates the transcription factor c-Myb, and miR-155 affects differentiation of CD4⁺ T lymphocytes into T_H1 cells, the development of regulatory T cells and the promotion of T cell-dependent tissue inflammation. miR-125b is up-regulated in human tonsil germinal
center B lymphocytes where it regulates Blimp-1 expression. miR-125b overexpression in hematopoietic stem cells causes myeloid leukemia in mice, and in humans miR-125b is over-expressed in hematological tumors such as acute lymphoblastic leukemia and acute myeloid leukemia.

Genome wide miRNA analyses have been performed in mouse lymphocyte subsets, in human tonsil derived B lymphocyte subsets and in total human CD4+ and CD8+ T cells, however little is known on the miRNAs that regulate differentiation of human T cell subsets. We performed a broad analysis of miRNA expression in seventeen different highly purified human lymphocyte subsets and identified specific miRNA signatures. In particular, we focused on naïve CD4+ T cells and found that miR-125b regulates a network of genes involved in CD4+ T cell differentiation including cytokine receptors, cytokines and transcription factors.
RESULTS

microRNA expression in primary human lymphocyte subsets

To assess miRNAs expression in human primary lymphocytes, seventeen T, B and natural killer (NK) cell subsets were purified by FACS according to established surface markers\textsuperscript{26-30} from peripheral blood mononuclear cells (PBMCs) of three to six healthy donors (Table 1a). The purified CD4\textsuperscript{+} T helper subsets were highly enriched for cells producing various marker cytokines (Table 1b), demonstrating the concordance between phenotype and effector functions. TaqMan miRNA Arrays were used to assess 664 miRNAs by RT-qPCR in these purified lymphocyte subsets. An unsupervised hierarchical clustering analysis of all the 242 miRNAs that were found expressed in the subsets analyzed provides an unbiased categorization of samples belonging to NK, T-CD4\textsuperscript{+}, T-CD8\textsuperscript{+} and B lymphocytes, which matches the phenotypic classification of subsets (Fig. 1a,b).

Previous reports have analyzed miRNA expression in five human lymphocyte subsets (naïve and memory CD4\textsuperscript{+} T cells, naïve CD8\textsuperscript{+} T cells, naïve B cells and NK cells)\textsuperscript{24,25}. Our analysis has identified most of the miRNAs (average of 78\%) already described in a particular cell subset by these studies. However, most miRNAs detected by our analysis have not been previously identified (Supplementary Fig. 1). Thus, our data provide a thorough quantitative and qualitative “Atlas” (Supplementary File 1) of the miRNome expression of most human lymphocyte subsets.

microRNA expression signatures of human lymphocytes

Next we searched for miRNAs that are preferentially expressed in selected human lymphocytes subsets, i.e., signatures (Supplementary Table 1). We therefore compared
the miRNomes of all seventeen cell subsets, selected those miRNAs that had more than 3
fold expression difference in a given subset (one way ANOVA p <0.01) and found that
twenty-nine miRNAs displayed subset specific expression (Fig. 1c). Within signatures,
we identified miRNAs that were not previously described as selectively expressed in any
human T lymphocyte subsets, such as miR-125b, miR-193b and miR-188-5p in naïve
CD4+ T cells, and miR-381 in CD4+ T_{H1} cells. In the case of a few miRNAs, such as
miR-146a and miR-21 in T_{reg} cells, we confirmed their specific expression^{31,32}.

We also looked for miRNAs that were differentially expressed among different
differentiation stages of a single lineage, such as T_{H1}, T_{H2}, and T_{H17} memory cells all
deriving from naïve CD4+ helper T cells and compared each of these four subset with the
other three. Twenty miRNAs were up-regulated and 5 down-regulated in naïve CD4+ T
cells compared to CD4+ memory subsets. Four miRNAs were found differentially
expressed in T_{H1}, two in T_{H17} and four were up-regulated in T_{H2} memory cells (Fig. 1d).

To validate the miRNA signatures identified in primary T lymphocytes ex vivo,
we followed their change in expression after activation in vitro. Naïve CD4+ T cells were
activated and miRNAs expression was monitored at different time points (up to 8 days).
Nineteen of the twenty miRNAs that were specifically highly expressed in resting naïve
CD4+ T cells were down-regulated upon activation, while there was an increase in the
expression of four out of five miRNAs that were highly expressed in resting CD4+
memory cells (Fig. 1e). We conclude that human CD4+ T lymphocyte subsets change
their specific miRNA signatures during differentiation from naïve to memory cells.

\textit{microRNA signatures in Homo sapiens and Mus musculus}
To assess the degree of conservation between the human and mouse immune system-specific miRNAs, we compared the human miRNomes and signatures we identified with those recently published in mouse lymphocyte subsets\(^{23}\). Considering the data currently available for mouse lymphocytes, we were able to compare six subsets (CD4\(^+\) naïve, CD4\(^+\) T\(_{H1}\), CD4\(^+\) T\(_{H17}\), CD4\(^+\) T\(_{H2}\), CD4\(^+\) T\(_{reg}\) and NK). Using the miRBase as reference, we compared the human and mouse lymphocyte-miRNomes by a BLAST search for similarities between mature miRNA sequences. An average 61\% of miRNAs expressed in the mouse lymphocyte subsets (Supplementary Fig. 2) were also expressed in the corresponding human lymphocytes. We focused our comparison on the forty-six miRNAs (Supplementary Table 2) that had been identified as signatures of the six human lymphocyte subsets (Fig. 1c) or as miRNA differentially expressed between naïve and effector CD4\(^+\) T cell subsets (Fig. 1d). Twelve of these forty-six miRNAs did not have a mouse homologue\(^{33}\) and twenty-two were not expressed in any of the six mouse lymphocyte subsets. Twelve of the forty-six miRNAs were expressed in mouse lymphocytes\(^{23}\) (Fig. 1f) but only six were expressed with the same profile in man and mouse lymphocyte subsets (Supplementary Fig. 3).

To rule out differences that might be due to different methods used in the different studies, we used RT-qPCR Taqman Assays to measure the expression of four miRNAs specific for human naïve CD4\(^+\) T cells (miR-125b, miR-188-5p, miR-193b and let-7c) in mouse naïve and memory CD4\(^+\) T cells. All four miRNAs were highly expressed in human naïve CD4\(^+\) T cells compared to human memory CD4\(^+\) T cells, while only let-7c was expressed in high amounts in naïve CD4\(^+\) T cells in mouse (Fig. 1g). We conclude there is a relatively good (61\%) conservation of the overall lymphocyte miRNomes, but a
poor concordance (6 out of 46) of the lymphocyte miRNA signatures between man and mouse.

**miR-125b targets genes in naïve CD4⁺ T cells**

Next we sought to identify the mRNA targets regulated by miRNAs in naïve CD4⁺ T cells that might be involved in the differentiation process toward effector-memory cells. We focused on miR-125b because it displayed the strongest differential expression between naïve CD4⁺ T cells and effector-memory lymphocyte subsets (Fig. 1c, Fig. 2a). We combined target prediction with gene expression analysis to identify those transcripts that exhibited an inverse relationship with the expression level of miR-125b (Anti-Correlated-Predicted-Target-mRNAs, ACEPTs) and found 72 ACEPTs for miR-125b (Fig. 2b and Supplementary Table 3). To assess the potential role of these 72 transcripts in T cell differentiation, we predicted their biological functions with Ingenuity Pathway Analysis, which maps bio-molecular networks based on known pathways, gene ontology and interactions with reliable data curation. The 72 miR-125b ACEPTs are significantly enriched for the biological function “Cell-mediated immune response” and the canonical pathway “T Helper Cell Differentiation” (Fig. 2c). Furthermore, miR-125b ACEPTs generate a network of 133 gene-products (Fig. 2d) that include several molecules involved in T cell differentiation such as cytokines, cytokine receptors and transcription regulators. In conclusion, the integrated analysis of gene expression and target prediction suggests that miR-125b regulates key genes involved in naïve CD4⁺ T cell differentiation.
miR-125b suppresses CD4+ T cell differentiation

We assessed the inhibitory effect of miR-125b on 14 genes belonging to the miR-125b ACEPT group and known to play a role in CD4+ T cell differentiation. The 3’ UTRs of these genes were cloned after a luciferase reporter gene and co-transfected into HEK-293T cells with a synthetic mature miR-125b double strand RNA or a scrambled control. miR-125b directly regulated four of the 14 UTRs, i.e., interferon-γ (IFNG), interleukin 2 receptor β (IL2RB), interleukin 10 receptor α (IL10RA) and Blimp-1 (PRDM1) (Fig. 3a). Deletion of the miR-125b responsive elements in these UTRs resulted in the abrogation of the inhibitory effect of miR-125b (Fig. 3b). Because miR-125b and its precursors are down-modulated in naïve CD4+ T cells after polyclonal activation in vitro (Fig. 1e and Supplementary Figure 5), we assessed the direct regulation of the four targets following miR-125b overexpression in primary cells. We assessed the expression levels of IFNG, IL2RB, IL10RA and PRDM1 transcripts by RT-qPCR three days after the transduction of naïve CD4+ T cells with lentiviral vectors expressing miR-125b (LV-125b) or a control lentivirus (LV-ctrl). Prolonged expression of miR-125b resulted in reduced levels (29 to 64%) of all four transcripts (Fig. 3c).

Compared with the human genes, the 3’UTR of the mouse IFNG and IL10RA genes do not have the predicted miR-125b responsive-elements, while the mouse IL2RB gene has one miR-125b responsive-elements compared to the three of the human gene. Such comparative analysis indicates that only the PRDM1 mouse gene has the predicted miR-125b responsive-elements (Fig. 3d and Supplementary Table 4). These data demonstrate that miR-125b directly regulates the naïve T cells expression of genes required for the differentiation of effector T lymphocytes. Most of miR-125b target genes
identified in man lack the miR-125b responsive-elements in the 3’UTR of the mouse homologues, correlating with the poor expression of miR-125b in naïve CD4⁺ T cells in mouse.

miR-125b enforces naïveness of CD4⁺ T cells

Because miR-125b directly regulates genes involved in the differentiation of naïve CD4⁺ T cells, we assessed the impact of miR-125b on the acquisition of the effector-memory phenotype. Naïve CD4⁺ T cells were activated by T cell receptor (TCR) engagement, transfected with synthetic miR-125b RNA (mimic-125b) and then stained with effector-memory marker mAbs. Three to five days after activation cells transfected with mimic-125b showed decreased surface expression of markers of Th1 and Th2 memory CD4⁺ T cells, such as CXCR3 and CCR4. In parallel, we observed an increased expression of the naïve CD45RA⁺ CD45RO⁻ phenotype, while control markers (CD4 and MHC class I) were not affected (Fig. 4a). In addition, mimic-125b decreased the expression of the two cytokine receptors (IL-2Rβ and IL-10Rα) that represent direct miR-125b targets (Fig. 4b). Analysis of CD4⁺ T cell proliferation with a Ki-67 mAb, which is expressed exclusively in proliferating cells, and using a CFSE dilution assay indicated that in vitro proliferation of naïve CD4⁺ T cells was not affected by transfection with mimic-125b (Fig. 4c). These results indicate that miR-125b expression contributes to the maintenance of naïveness in primary human T cells and suggest that its down-modulation is associated with the acquisition of effector-memory phenotypes by CD4⁺ T cells.

miR-125b regulates acquisition of effector function
To assess whether prolonged expression of miR-125b affected the acquisition of effector function, naïve CD4\(^+\) T cells were activated and transduced with lentiviral vectors encoding pre-miR-125b (LV-125b) or a control lentivirus (LV-ctrl). After 8 days of culture, cells were tested for effector function after stimulation for 6 hours with phorbol 12-myristate 13-acetate (PMA) plus ionomycin. Naïve CD4\(^+\) cells activated in the presence of ectopic miR-125b displayed a reduced effector function as demonstrated by the decreased production of IFN-\(\gamma\) and IL-13 detected both intracellularly (Fig. 5a,b) and in the culture supernatants (data not shown), whereas cells transduced with the control lentivirus showed effector function similar to untransduced controls. Comparable results were obtained using a lentivirus vector expressing a mutated miR-125b, as well as mimic-125b (Supplementary Fig. 4), miR-125b expression in cells transduced with LV-125b was similar to that observed in untransduced primary naïve cells (Fig. 5c), indicating that ectopic expression of miR-125b was within the physiological range. These results strongly suggest that miR-125b plays a key role in preventing the acquisition of effector functions by primary CD4\(^+\) T cells.
Here we perform an analysis of the miRNomes in seventeen major human lymphocyte subsets and show that the specific miRNAs signatures that mark these subsets have a poor concordance between human and mouse lymphocytes. Identification of miRNA target genes shows that miR-125b, one of the miRNA that is highly specific for naïve CD4+ T cells, directly regulates a set of genes important to the differentiation into effector-memory lymphocytes and its prolonged expression during differentiation reduces the acquisition of effector function.

miRNomes provide a molecular categorization of different lymphocyte subsets that is coherent with their phenotypic characterization and suggests that selected miRNAs play a role in regulating the differentiation of these cell subsets. The comparison of the subset-specific miRNomes identified in this study with the published human miRNomes revealed that our approach identified up to 80% of the published miRNAs and identified additional miRNAs, possibly due to the sensitivity of the approach. Because TaqMan miRNA Arrays require only a few nanograms of RNA, this approach is suitable for very rare lymphocyte subsets, allowing us to obtained the widest available microRNA profile of human lymphocyte subsets.

The comparison of miRNomes obtained in various studies and from different species can be challenging, because the nomenclatures used in different technological platforms are not always aligned. To address this issue we first created a database that incorporates and integrates the different nomenclatures of the various miRBase versions. Secondly, because deep sequencing technology often identifies microRNAs with codes that are different from standard miRBase nomenclature, we used the mouse miRNA
genomic coordinates and mature sequences to translate the deep sequencing data into miRBase-compliant nomenclature. Thus, comparison of the miRNA signatures between mouse and man lymphoid subsets was possible and revealed a surprisingly low 13% concordance. In particular, three (miR-193b, miR-188-5p and miR-125b) of the four miRNAs that constitute a signature of naïveness in human CD4⁺ T cells are almost absent in mouse cells.

The identification of miRNA target genes is challenging due to the fact that miRNAs bind to target mRNAs by incomplete sequence pairing over a short sequence. Several in silico, molecular and biochemical approaches have been used in the last few years to identify miRNA targets, each one with strengths and weaknesses. Our approach is extremely helpful in restricting the number of putative target genes that can be predicted in silico. Indeed, using the miR-125b ACEPT gene list, we restricted our target gene search from the 2058 predicted target genes to 72 miR-125b ACEPT genes. Cloning of the whole 3’UTRs and dual-luciferase assay revealed that 4 out of 14 of these 72 potential targets were directly regulated by miR-125b.

Our findings suggest that the naïve state is, at least in part, actively preserved by miR-125b that reduces the expression of important genes below the levels required to confer effector-memory identity to a naïve cells. The TH1 cytokine IFN-γ represents a highly relevant target, because it is the most abundant effector cytokine in human memory T cells and it is not secreted by naïve T cells. The IL-2Rβ chain is also critical for memory cell identity, because it mediates IL-15-dependent homeostatic self-renewal of memory T cells, while naïve T cells lack IL-2Rβ, do not respond to IL-15 and their homeostasis depends on IL-7 and TCR tickling induced by self-MHC. Blimp-1 is
known to promote terminal T cell differentiation\textsuperscript{39}. miR-125b alone cannot account for all the complex series of changes in gene expression and functions that mark differentiation of naïve to memory T cells, and it is likely the other miRNA signatures of naïve CD4\textsuperscript{+} T cells (miR-188-5p, miR-193b and let-7c) play a role. However, the importance of the miR-125b targets outline its relevance in the enforcement of the naïve state of CD4\textsuperscript{+} T cells.

Although mouse and man mostly utilize these same genes for the differentiation of naïve to memory T cells, we found that three (IFN-\(\gamma\), IL10R\(\alpha\) and IL2R\(\beta\)) of the four genes that are the targets of miR-125b in man, do not have the miR-125b responsive elements in their 3’UTR in mouse. This findings, together with the observation that mouse naïve T cells express very little miR-125b, suggest a dichotomy between the genes key to naïve T cell differentiation and their regulatory miRNAs between man and mouse, and underscore the necessity of studying the role of miRNAs in human lymphocyte differentiation primarily with human cells.

From a translational medicine point of view, the importance of deciphering the miRNome in human T lymphocyte is twofold. First, identification of miRNA signatures and definition of their target genes that specifically regulate differentiation of the various T cell subsets may provide novel therapeutic targets for immune-mediated diseases and give a better definition of subsets. Second, because miRNAs are released in a stable form outside cells\textsuperscript{40}, it is possible that lymphocyte subset miRNA signatures circulate in the blood in a cell-free state, and could therefore be used as sensitive biomarkers of different T cell subsets that are active in distant tissues.
DATABASE ACCESSION NUMBERS

Raw and normalized gene expression data are available for download in the GEO format with accession number GSE22880.

ACKNOWLEDGEMENTS

We thank P. Dellabona, P. De Candia and S. Monticelli for helpful discussions. L.W. is supported by a grant from the National Research Fund of Luxembourg (PDR-08-040). Part of this study was supported by a grant from the Fondazione Cariplo (Milano, Italy).

CFI

All authors declare that they have no competing financial interests with this work.

AUTHOR CONTRIBUTIONS

R.L.R, G.R. and L.W. designed and did experiments, analyzed data and wrote the manuscript; M.M. and M.C.C. did FACS experiments; R.J.P.B. performed bioinformatic analyses; M.C. did statistical analyses; R.S.B., S.C., S.M., P.G., F.M., D.M. V. P. and A.R. did experiments and analyzed data; E.T. provided conceptual advice; R.D.F. discussed the results, provided conceptual advice and commented on the manuscript; J.G. designed and supervised research and wrote the manuscript; S.A. and M.P. designed the study, supervised research and wrote the manuscript.
Table 1 Purification of human primary lymphocyte subsets. (a) Seventeen human lymphocyte subsets were isolated from Peripheral Blood Lymphocytes. Purity is that achieved after sorting; Sorting phenotype indicates surface marker combinations used for cell sorting. (b) Cytokines produced by the purified T helper cell subsets following polyclonal stimulation with PMA and ionomycin. Cells co-producing IFN-γ and IL-4 or IFN-γ and IL-17 were very low (< 1%) or undetectable (data not shown).

Figure 1 microRNA expression describes and discriminates among different lymphocyte subsets. (a) Unsupervised hierarchical clusters of the miRNAs detected for each cell subset analyzed. 242 subset-specific miRNAs were selected using one-way Anova analysis (p-value 0.001). Yellow to blue gradient indicates expression values ($\Delta$CT log transformed and mean centered). The bracketed letters in the sample names indicates the individual donor. (b) Cell-lineages segregation and colour codes of various lymphocyte-subsets as in panel a (c) Heat map of average expression values (z-scores of $\Delta$CT) for mature miRNAs across all 17 lymphocyte subsets profiled. (d) z-scores heat map of signature miRNAs for naïve CD4$^+$, T$_{H1}$, T$_{H17}$ and T$_{H2}$ cell subsets. Differentially expressed miRNAs were selected with the comparative CT method ($\Delta$ΔCT) and filtered for statistical significance Anova test (p < 0.05). Each one of the four subsets was in turn considered as reference and compared with the remaining three subsets. (e) Heat map of relative quantities (RQ) for 25 miRNAs assessed by RT-qPCR in naïve CD4$^+$ T cells activated with anti-CD3 and anti-CD28 mAbs at the indicated time points compared to time zero. (f) Expression of 46 miRNAs representing the human lymphocytes signatures.
of the six lymphocyte subsets ( naïve CD4\(^+\) T cells, T\(_{H1}\), T\(_{H17}\), T\(_{H2}\) and NK cells) for which mouse data are published\(^23\). (g) Expression of miR-125b, miR-188-5p, mir-193b and let-7c assessed by TaqMan assays in naïve and memory CD4\(^+\) T cells purified from human PBMC and mouse spleen.

**Figure 2** Identification of putative miR-125b target genes in naïve CD4\(^+\) T cells and description of their functional relationships. (a) RT-qPCR data of miR-125b expression in lymphocyte subsets is shown as mean expression values. (b) Venn diagram of CD4\(^+\) naïve and memory cells transcripts and miR-125b predicted targets (TargetScan 5.1). 857 are EPTs (Expressed Predicted Targets -in red-), Pie chart shows the distribution of EPT genes according to their memory/naïve (M/N) expression ratio. (c) Enrichment of the “T Helper Cell Differentiation” canonical pathway genes, according to IPA, in the EPT (M/N>1.5) group. Control groups were the expressed genes (M/N>1.5), EPT (M/N<1.5) and the predicted targets. Significance is plotted as –\(\log_2\)(p-values). (d) Molecular network of functional relationships extracted from IPA knowledge base for the EPT with a M/N>1.5. The graph shows direct interactions generated by genes that are “network eligible” according to IPA (dark blue); Molecules in the network that are predicted to be miR-125b targets are in light blue; In white other network molecules.

**Figure 3** miR-125b directly regulates genes key to T cell differentiation. (a) Dual luciferase assay of 14 genes predicted to be regulated by miR-125b. The luciferase constructs was transfected into HEK-293T cells together with synthetic mature miR-125b (black bars) or a synthetic scrambled control (white bars). An empty vector and a non
miR-125b target transcript (GAPDH) were used as negative controls while a vector with a tandem stretch of six miR-125b target sites (Sponges) was the positive control. The dashed line represents the threshold (75%) of luciferase activity. The values represent average ± SEM (n=18). T-test results are indicated by (***) P<0.001 relative to the scrambled transfected controls. (b) miR-125b miRNA responsive element was deleted in the 3′UTR of *IFNG*, *IL10RA*, *IL2RB* and *PRDM1*, and luciferase activity was assessed as described above. The values represent average ± SEM (n=18) Anova test significance is indicated by (***) P<0.001. (c) Expression levels of *IFNG*, *IL10RA*, *IL2RB* and *PRDM1* transcripts were measured by RT-qPCR in activated naïve CD4+ T cells transduced with lentiviral vectors encoding miR-125b (LV-125b) or mock control (LV-ctrl). The values represent average ± SEM of 3 independent experiments. (d) Scheme of miR-125b responsive elements (colour coded) predicted by Targetscan5.1 in human and mouse *INFG*, *IL10RA*, *IL2RB* and *PRDM1* transcripts.

**Figure 4** miR-125b preserves naïve state of CD4+ T cells. Purified peripheral blood naïve CD4+ T cells were stimulated with anti-CD3 and allo-PBMCs and nucleofected with mimic-125b or scrambled control. (a) Cells were stained at day 3 for CXCR3 and CCR4 or at day 5 for CD45RA, CD45RO, CD4 and MHCI surface expression by flow cytometry. (b) Cells were analyzed after three days for IL10Rα, and IL2Rβ surface expression by flow cytometry. (c) Purified, unlabeled or CFSE-labeled naïve CD4+ T cells were stimulated with anti-CD3 and allo-PBMCs and nucleofected with mimic-125b or scrambled control, proliferation was assessed respectively by Ki67 staining or CFSE dilution at day 5. Numbers indicate percentages. Similar results were obtained in three
Figure 5 miR-125b regulates naïve CD4⁺ T cell differentiation. (a) and (b) Purified naïve CD4⁺ T cells were stimulated with anti-CD3 and anti-CD28 antibodies and transduced with lentiviral vectors encoding miR-125b (LV-125b), or mock control (LV-ctrl). After seven days, cells were re-activated with PMA and Ionomycin and the acquisition of effector function was assessed as production of IFN-γ and IL13 by intracellular staining after 6 hours. Mean and statistical significance of six independent experiments are shown in panel b. (c) miR-125b expression levels measured by RT-qPCR in naïve CD4⁺ T cells stimulated for seven days with anti-CD3 and anti-CD-28 antibodies, which have been transduced with control (LV-ctrl) or lentiviral vectors encoding miR-125b (LV-125b). As control we used untransduced primary naïve CD4⁺ T cells and CD4⁺ T_{H1} memory cells. The mean of 3 independent experiments is shown.
METHODS

Purification of human lymphocyte subsets and naïve T cell differentiation experiments.

Buffy-coat blood of healthy donors was obtained from the IRCCS Policlinico Ospedale Maggiore in Milano, Italy and PBMCs were isolated by Ficoll centrifugation. Lymphocyte subsets were purified by cell sorting on a FACS Aria (BD) using different combinations of surface markers (Table 1). The ethical committee of IRCCS Policlinico Ospedale Maggiore in Milano (Italy) approved the use of PBMCs of healthy donors for research purposes and informed consent was obtained from all the subjects involved in this study. For in vitro differentiation to effector cells, resting naïve CD4\(^+\) T lymphocytes were purified (> 98%) by negative selection with magnetic beads and activated at 1x10^5 cells/well on 96 well MaxiSorp plates (Nunc) coated with anti-CD3 (0.1 \(\mu\)g/ml) and anti-CD28 (6 \(\mu\)g/ml) antibodies (BD) and IL-2 at 20 IU/ml (Novartis). After seven days, cells were stimulated for 6 hrs with PMA-Ionomycin and effector function was assessed as IFN-\(\gamma\) production by intracellular staining with Pb-conjugated anti-IFN-\(\gamma\) antibody (BD). IL-13 production was assessed with PE-conjugated anti-IL13 antibody (BD)

RNA isolation and miRNA-mRNA expression profiling.

Total RNA was isolated using the mirVana Kit (Ambion). The Megaplex protocol with pre-amplification was performed. Pre-amplified RT product was amplified using TaqMan Low Density Arrays (Applied Biosystems). Gene expression was performed on naïve CD4\(^+\) and memory T cells with Illumina direct hybridization assays: cRNA was generated according to the Illumina protocol (Ambion); Hybridization and scanning were
performed on an Illumina iScan System and data were processed with BeadStudio v.3. Both raw and normalized gene expression data are available for download in the GEO format with accession number GSE22880.

Data treatment, statistics and bioinformatics
The median of three endogenous controls was used as computed control and ΔCT were calculated. Raw CT values are available in the supplementary materials. MiRNA expressed in at least two thirds of samples analyzed were selected with a 1-way ANOVA test. miRNAs signatures for specific cell subsets were determined as detailed in the supplemental experimental procedures. TargetScan 5.1 was used for miR-125b target predictions; expressed predicted targets were selected comparing the predicted targets with the genes expressed in naïve and-or memory CD4⁺T cells from gene expression experiments. Pathway-function analysis was performed with IPA and statistical relevance was determined with the built-in Fisher exact test.

Molecular biology
miRNA expression levels on specific human and mouse populations was made using Taqman MiRNA assays (Applied Biosystems) according to the manufacturer protocol. miR-125b direct target expression levels were assessed by TaqMan Gene Expression assays (Applied Biosystems). 3’UTRs of 14 human miR-125b target genes were amplified from Human Testes or Human Fetal Brain cDNA, transferred to psiCHECK-2 vector (Promega). The sponge control vector was designed as a repeat of six miR-125b
responsive elements (MRE). Mutated 3’UTRs were obtained deleting the miR-125b MRE in the 3’ UTR by PCR mutagenesis, except for IL2Rβ’s construct that was synthesized (Geneart AG). Precursor miR-125b encoding lentiviral vector (pLV-125b) was purchased by SBI. Primers sequences are described in the supplementary experimental procedures.

**Dual luciferase assay**

HEK293T cells were cotransfected with 10 ng of each psiCHECK-2 construct along with 20nM miR-125b duplexes or scrambled control (Qiagen) using Lipofectamine2000 (Invitrogen). After 24 hours, the cells were lysed and Firefly and Renilla Luciferase activities were measured with the Dual Luc Reporter System (Promega). The whole 3’UTRs were cloned in psiCHECK-2 immediately downstream the Renilla luciferase gene. Results were expressed as the ratio of Luciferases activities.

**Naïve CD4⁺ T cells nucleofection**

Purified CD4⁺ naïve T Cells, unloaded or pre-loaded with CFSE, were nucleofected with miR-125b duplexes or scrambled control (Dharmacon) using T Cell Nucleofector kit (VAPA-1002 Lonza). Cells were plated in complete medium and stimulated with soluble αCD3 0.1 μg/ml (clone TR66) and allo-PBMCs in a 96 well U-bottom plate in the presence of IL2 100 U/ml, αIL4, αIFN-γ and αIL12 (R&D) antibodies at 2 μg/ml. Differentiation and proliferation markers and miR-125b direct targets were analyzed after surface staining using the following fluorescence-labeled antibodies from BD Biosciences: APC-CD45RO (clone UCHL1), PECy7-CD45RA (clone-H1100), PE-CD122 (clone MIKβ3), PB-CD4 (clone RPA-T4), PECy5-CD183 (clone 1C6/CXCR3),
intracellular staining was done according to the manufacturer’s protocol for Ki67 using PE-Ki67 (clone B56). A FACSCantoII and FlowJo887 software were used for data acquisition and analysis.

**Lentivirus production and transduction**

Lentiviral particles were produced according to the standard protocol indicated in the SBI User Manual instructions. Naïve CD4$^+$ T cells were activated as described above and simultaneously transduced with either pLV-ctrl vector or pLV-125b at MOI: 1x10$^7$ TU/ml in the presence of polybrene (8μg/ml). Cells were detached on day 3 and transduction efficiency measured by the percentage of GFP positive cells by flow cytometry. Transduced cells were then replated on uncoated wells and cultured as described above.
REFERENCES


## Table 1

### a

<table>
<thead>
<tr>
<th>Lymphocyte subsets</th>
<th>Number of donors</th>
<th>Purity (%)</th>
<th>Sorting phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD4(^+) naïve</td>
<td>4</td>
<td>99.8 ±0.1</td>
<td>CD4+, CCR7+, CD45RA+, CD45RO-</td>
</tr>
<tr>
<td>CD4(^+) T(_{H1})</td>
<td>6</td>
<td>99.9 ±0.05</td>
<td>CD4+, CXCR3+, CCR6-, CD161-</td>
</tr>
<tr>
<td>CD4(^+) T(_{H17})</td>
<td>4</td>
<td>99.1 ±1</td>
<td>CD4+, CCR6+, CD161+, CXCR3-</td>
</tr>
<tr>
<td>CD4(^+) T(_{H2})</td>
<td>4</td>
<td>99.7 ±0.3</td>
<td>CD4+, CRTH2+, CXCR3-</td>
</tr>
<tr>
<td>CD4(^+) Treg</td>
<td>3</td>
<td>99.0 ±0.8</td>
<td>CD4+, CD127dim, CD25+</td>
</tr>
<tr>
<td>CD4(^+) memory</td>
<td>3</td>
<td>99.8 ±0.2</td>
<td>CD4+, CD45RA-, CD45RO+</td>
</tr>
<tr>
<td>CD4(^+) T(_{EM})</td>
<td>3</td>
<td>98.4 ±2.8</td>
<td>CD4+, CCR7+, CD45RA-, CD45RO+</td>
</tr>
<tr>
<td>CD4(^+) T(_{EM})</td>
<td>3</td>
<td>95.4 ±5.5</td>
<td>CD4+, CCR7-, CD45RA-, CD45RO+</td>
</tr>
<tr>
<td>CD4(^+) T(_{EMRA})</td>
<td>3</td>
<td>96.5 ±4</td>
<td>CD4+, CCR7-, CD45RA+, CD45RO-</td>
</tr>
<tr>
<td>CD8(^+) naïve</td>
<td>3</td>
<td>99.3 ±0.2</td>
<td>CD8+, CCR7+, CD45RA+, CD45RO-</td>
</tr>
<tr>
<td>CD8(^+) T(_{CM})</td>
<td>3</td>
<td>98.3 ±0.8</td>
<td>CD8+, CCR7-, CD45RA-, CD45RO+</td>
</tr>
<tr>
<td>CD8(^+) T(_{EM})</td>
<td>3</td>
<td>96.8 ±0.9</td>
<td>CD8+, CCR7-, CD45RA-, CD45RO+</td>
</tr>
<tr>
<td>CD8(^+) T(_{EMRA})</td>
<td>3</td>
<td>98.1 ±0.3</td>
<td>CD8+, CD45RA+, CCR7-, CD45RO-</td>
</tr>
<tr>
<td>NK cells</td>
<td>3</td>
<td>99.2 ±0.3</td>
<td>CD56+, CD3-</td>
</tr>
<tr>
<td>CD5(^+) B cells</td>
<td>3</td>
<td>99.1 ±0.8</td>
<td>CD19+, CD5+</td>
</tr>
<tr>
<td>naïve B cells</td>
<td>3</td>
<td>99.9 ±0.1</td>
<td>CD19+, CD5-, CD27-</td>
</tr>
<tr>
<td>memory B cells</td>
<td>3</td>
<td>99.1 ±0.8</td>
<td>CD19+, CD5-, CD27+</td>
</tr>
</tbody>
</table>

### b

<table>
<thead>
<tr>
<th>CYTOKINES (%)</th>
<th>IFN-γ</th>
<th>IL-17</th>
<th>IL-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>T(_{H1})</td>
<td>50.5</td>
<td>0.9</td>
<td>0.6</td>
</tr>
<tr>
<td>T(_{H17})</td>
<td>4.5</td>
<td>13.1</td>
<td>0.6</td>
</tr>
<tr>
<td>T(_{H2})</td>
<td>0.6</td>
<td>0.6</td>
<td>57</td>
</tr>
</tbody>
</table>
Figure 1

(a) Heatmap showing gene expression levels with z-scores. High expression is indicated in red, low in blue.

(b) Trend plot showing changes in gene expression over time after activation.

(c) Scatter plot comparing miRNA expression levels in different cell types.

(d) Expression levels of miRNAs in CD4+ naive and TEM cells.

(e) Stability of miRNA expression over time.

(f) Venn diagram showing overlap in miRNA expression between different cell populations.

(g) Bar graph comparing miRNA expression in human and mouse cells.
Figure 2

a

![Graph showing miR-125b expression levels in different cell populations.]

Naive | Memory | miR-125b PTs
---|---|---
2045 | 6056 | 222
214 | 617 | 26
1201 | | |

b

![Venn diagram illustrating the overlap of predicted targets (miR-125b PTs) in Naive and Memory cell populations.]

<table>
<thead>
<tr>
<th>Naive only</th>
<th>M/N &gt; 1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>214</td>
<td>72</td>
</tr>
<tr>
<td>1.5 &gt; M/N &gt; 0</td>
<td></td>
</tr>
<tr>
<td>6056</td>
<td></td>
</tr>
<tr>
<td>617</td>
<td></td>
</tr>
<tr>
<td>2045</td>
<td></td>
</tr>
<tr>
<td>222</td>
<td></td>
</tr>
<tr>
<td>1201</td>
<td></td>
</tr>
</tbody>
</table>

EPT (M/N > 1.5) | EPT (M/N < 1.5) | Predicted target (unexpressed)
---|---|---
na | na | |

M/N > 1.5Naive only

$\log (P$ value $)$

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>na</td>
<td>na</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TRANSCRIPTION FACTOR

Predicted targets

Cytokine

Other

Transmembrane receptor

Unknown

Extracellular space

Plasma membrane

Cytoplasm

Nucleus

Unknown

ACEPTs

Predicted targets

Transmembrane receptor

Other

Cytokine

Figure 2

100 200 300 400 500

CD4 $^+$ naive

CD4 $^+$ TH1

CD4 $^+$ TH17

CD4 $^+$ TH2

CD4 $^+$ Treg

CD4 $^+$ memory

CD4 $^+$ TEM

CD4 $^+$ TCM

CD4 $^+$ TEMRA

CD8 $^+$ naive

CD8 $^+$ TEM

CD8 $^+$ TEM

CD8 $^+$ TEMRA

CD5 $^+$ B

naive B

memory B

NK
Figure 3

(a) Luciferase activity (relative)

(b) Luciferase activity (relative)

(c) Expression Units

(d) Expression Units
Figure 5

(a) Flow cytometry plots showing the expression of IFN-γ and IL-13 in different groups:
- **Unt**: Untreated
- **LV-ctrl**: LV-ctrl
- **LV-125b**: LV-125b

(b) Bar graphs illustrating the expression levels of IFN-γ and IL-13:
- Unt
- LV-ctrl
- LV-125b

(c) Relative expression of miR-125b in different conditions:
- Naive
- Th1
- LV-ctrl
- LV-125b