Memory T and memory B cells share a transcriptional program of self-renewal with long-term hematopoietic stem cells

Chance John Luckey*†, Deepta Bhattacharya*†‡, Ananda W. Goldrath*†§, Irving L. Weissman‡, Christophe Benoist*¶, and Diane Mathis*¶

*Joslin Diabetes Center; Departments of Pathology and Medicine, Brigham and Women's Hospital, Harvard Medical School, 1 Joslin Place, Boston, MA 02215; and ‡Stanford Institute for Stem Cell Biology and Regenerative Medicine, Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305-5323

Contributed by Diane Mathis, December 23, 2005

The only cells of the hematopoietic system that undergo self-renewal for the lifetime of the organism are long-term hematopoietic stem cells and memory T and B cells. To determine whether there is a shared transcriptional program among these self-renewing populations, we first compared the gene-expression profiles of naïve, effector and memory CD8$^+$ T cells with those of long-term hematopoietic stem cells, short-term hematopoietic stem cells, and lineage-committed progenitors. Transcripts augmented in memory CD8$^+$ T cells relative to naïve and effector T cells were selectively enriched in long-term hematopoietic stem cells and were progressively lost in their short-term and lineage-committed counterparts. Furthermore, transcripts selectively decreased in memory CD8$^+$ T cells were selectively down-regulated in long-term hematopoietic stem cells and progressively increased with differentiation. To confirm that this pattern was a general property of immunologic memory, we turned to independently generated gene expression profiles of memory, naïve, germinal center, and plasma B cells. Once again, memory-enriched and -depleted transcripts were also appropriately augmented and diminished in long-term hematopoietic stem cells, and their expression correlated with progressive loss of self-renewal function. Thus, there appears to be a common signature of both up- and down-regulated transcripts shared between memory T cells, memory B cells, and long-term hematopoietic stem cells. This signature was not consistently enriched in neural or embryonic stem cell populations and, therefore, appears to be restricted to the hematopoeitic system. These observations provide evidence that the shared phenotype of self-renewal in the hematopoietic system is linked at the molecular level.

Self-renewal is a process by which a daughter cell that maintains the same properties as its parent is generated. The best-studied self-renewing cells are long-term hematopoietic stem cells (Lt-HSC), which maintain themselves as a population for the lifetime of the organism. However, self-renewal within the hematopoietic system is not limited to stem cells, because antigen-specific memory B and T cells have also been observed to self-renew in perpetuity. Although this phenotypic similarity has been noted previously (1–3), there is to date no information on whether these cells use the same molecular pathways for self-renewal. Although the extracellular signals involved in cellular homeostasis likely differ between memory and stem cells, we hypothesized that these external cues converge on some of the common cell-intrinsic mediators involved in self-renewal, perhaps through the reactivation of genetic programs used by Lt-HSC.

Adult Lt-HSC are multipotent cells capable of both lifelong self-renewal and differentiation into the various mature cellular components of blood (4). Differentiation of Lt-HSC leads to the formation of short-term hematopoietic stem cells (St-HSC). Although St-HSC retain full hematopoietic differentiation potential, they have a more limited, "short-term," self-renewal potential. St-HSC subsequently differentiate into lineage-committed precursors (LCP) of either the myeloid or lymphoid lineages. Further differentiation of LCP is restricted to their respective lineage, and they are incapable of self-renewal. The inability to undergo self-renewal holds true for all subsequent downstream precursor populations as well as for the majority of mature blood cells. Thus, the self-renewal of Lt-HSC is required for sustained hematopoiesis over the course of an organism's life.

Memory T and B cells are mature blood cells that reacquire the ability to undergo long-term self-renewal and are the product of a carefully controlled process of differentiation in response to immunostimulation, such as infection by pathogens (1–3, 5, 6). Before infection, antigen-inexperienced, or "naïve," cells of a particular specificity exist at very low frequencies and rarely, if ever, divide (7–9). Upon antigenic exposure, naïve cells capable of recognizing one of the pathogen's components undergo a process of rapid clonal expansion and differentiation. For T cells, this process leads to the generation of effector cells that have acquired the functional capacity to rapidly combat foreign pathogens. Effector T cells undergo a dramatic contraction in numbers after pathogen clearance, with 90–95% of them succumbing to apoptosis within weeks after the initial infection (2, 5). However, a subset of the antigen-specific cells persists long after antigen exposure and constitutes the memory T cell compartment.

For B cells, the early thymus-dependent responses to antigenic challenge lead to the formation of rapidly proliferating, short-lived, antibody-secreting plasma cells and germinal center B cells, which undergo somatic hypermutation and Ig isotype switching. Similar to effector T cells, the vast majority of these two cell types is eliminated through apoptosis (10, 11). The surviving antigen-specific B cells comprise two separate memory compartments: the long-lived antibody-secreting plasma cell and the self-renewing memory B cell. The antibody-secreting plasma B cells are completely quiescent and secrete antigen-specific Ig indefinitely, irrespective of antigen reexposure (12). In contrast, self-renewing memory B cells proliferate slowly and rapidly respond to antigen reexposure by differentiating into both plasma and germinal center B cells in another round of affinity maturation (10, 13).

Memory lymphocytes respond more robustly than their naïve counterparts to antigenic challenge. This ability to respond, combined with their increased frequency and self-renewal, ensures that reexposure to a particular pathogen leads to rapid and vigorous

---

Conflict of interest statement: No conflicts declared.

Abbreviations: LCP, lineage-committed precursors; Lt-HSC, long-term hematopoietic stem cells; St-HSC, short-term hematopoietic stem cells.

Data deposition: All data sets have been deposited in the Gene Expression Omnibus, www.ncbi.nlm.nih.gov/geo (accession nos. GSE 4142 and GSM 94871–94882).

†C.J.L., D.B., and A.W.G. contributed equally to this work.

§Present address: Division of Biology Section of Molecular Biology, University of California at San Diego, 9500 Gilman Drive, San Diego, CA 92093-0377.

¶To whom correspondence should be addressed. E-mail: cbdm@joslin.harvard.edu.

www.pnas.org/cgi/doi/10.1073/pnas.0511137103

cellular and humoral responses. Thus, memory B and T cells, like long-term hematopoietic stem cells, retain the ability to further differentiate when called on while maintaining themselves through a process of self-renewal. We hypothesized that these similarities would be reflected in a common transcriptional profile.

Results

To broadly compare the gene expression profiles of HSC and T cell populations, we used Affymetrix GeneChip technology. For the initial analysis, we first used the Genechip data generated by Ivanova *et al.* (14) in which phenotypically well characterized and functionally defined mature HSC populations were compared with both St-HSC and LCP. As a source of monoclonal populations of naïve, effector, and memory T cells, we used the OTI T cell receptor transgenic mice because this strain greatly facilitates the functional definition and purification of the $CD8^+$ T cells at various stages of differentiation. We purified naïve, effector, and memory $CD8^+$ T cells as described in ref. 15. Each population expressed the expected markers (e.g., cell-surface) and exhibited diagnostic functional activities. In particular, we demonstrated that the memory T cells generated in our system self-renew by following them in a cohort of mice over the course of several months. OTI memory T cells maintained constant cell numbers, cell-surface phenotypes, and cytokine secretion profiles (C.J.L., A.W.G., C.B., and D.M., unpublished results), consistent with the published reports demonstrating that $CD8^+$ memory T cells homeostatically self-renew over the lifetime of an individual (16, 17).

Having established functionally defined cell populations, we grouped the raw data (.cel files) and collectively preprocessed them with the "rma" method in the AFFYLMGUI statistical analysis package (detailed in *Methods*). Diagnostic transcripts behaved as expected, confirming that the data accurately depict the gene-expression profiles of naïve, effector, and memory cells. For example, the microarray-measured levels of CD44, IL-7 receptor, IL-15 receptor, IL-2 receptor, Bcl-2, Bcl-X, granzyme A, granzyme B, and IFN-γ transcripts all showed the same order of expression in naïve, effector, and memory cells as has been reported for their mRNA and/or protein in previous studies (ref. 15; see also Fig. 5*A*, which is published as supporting information on the PNAS web site). Furthermore, the profiles were remarkably similar to those independently generated by using a different T cell receptor-transgenic mouse/viral infection system, providing an important external validation of the experimental approach (ref. 18; Fig. 5*B*). Finally, the GeneChip data correctly predicted differential expression as measured by quantitative PCR (Q-PCR) $\approx$90% of the time, although often dramatically underestimating Q-PCR-estimated fold changes (Fig. 5*C*).

To identify transcripts augmented in memory cells, we first compared the gene expression profiles of memory and naïve cells by displaying them as a function of *P* value versus fold change (Fig. 1*A*). The *x* axis represents the $\log_2$ (fold change), so that transcripts that are equally expressed in the two cell populations are zero and fall on the gray midline. A transcript whose expression is relatively enriched in a given population shows up away from the midline, toward the side where its expression is highest. The *y* axis represents the false discovery rate-corrected *P* value, with those transcripts with the lowest *P* values having the highest likelihood of being truly different in the two cell populations being compared. Thus, those transcripts whose expression is the lowest and farthest from midline are the most likely enriched in the population toward which they are skewed. The transcripts whose levels were most increased in memory cells were selected, and their expression values in memory versus effector cells were secondarily plotted (Fig. 1*B*). Thereby, a set of 98 transcripts enriched in memory $CD8^+$ T cells relative to both naïve and effector $CD8^+$ T cells was delineated. Analysis of these enriched transcripts within the various stem cell populations revealed a preferential representation in Lt-HSC-enriched genes

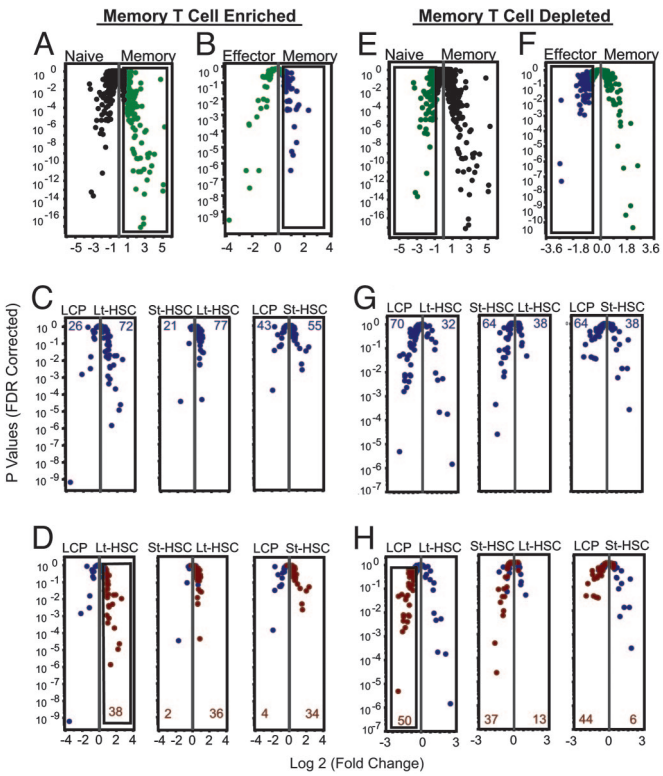

Fig. 1. Memory $CD8^+$ T cell-enriched transcripts are also enriched in Lt-HSC, whereas memory $CD8^+$ T cell-depleted transcripts are also depleted in Lt-HSC. Volcano plots comparing the relative gene expression in the two cell populations listed above each plot are shown. For memory-enriched transcripts, all 12,422 transcripts on the MgU74v2 A chip are shown for the memory vs. naïve comparison (*A*). Transcripts whose expression was relatively enriched in memory cells (unlogged fold change $\geq$1.4) are shown in green and were used for subsequent analysis. This criterion was based on within-replicate analyses providing empirical estimation of experimental noise. (*B*) Relative expression in the memory vs. effector cell comparison for those transcripts selected in *A*. Transcripts whose fold change is $\geq$1.4 are shown in blue and were used for subsequent analysis. The transcripts meeting criteria in both *A* and *B* were then plotted for their relative expression in *C* as Lt-HSC vs. LCP, Lt-HSC vs. St-HSC, and St-HSC vs. LCP comparisons. In each case, the number of transcripts whose $\log_2$ (fold change) is greater or less than zero is shown at the top of the plot. Those transcripts whose expression in Lt-HSC vs. LCP was $>$1.4 were then highlighted in red in *D*. In each case, the numbers of red transcripts whose $\log_2$ (fold change) is greater or less than zero is shown at the bottom of the plot. For memory-depleted transcripts, the 98 transcripts whose expression was relatively depleted (fold change of $\leq$1.4) in memory T cells relative to naïve (*E*) and effector (*F*) T cells are plotted for their relative expression in Lt-HSC vs. LCP, Lt-HSC vs. St-HSC, and St-HSC vs. LCP (*G*). The number of transcripts whose $\log_2$ (fold change) is greater or less than zero is shown at the top of the plot. Those transcripts whose expression in Lt-HSC vs. LCP was <-1.4 were then highlighted in red (*H*), and their number greater or less than zero is shown at the bottom of the plot.

compared with either LCP or St-HSC, both in numbers (73% and 78% respectively) and *P* values (Fig. 1*C*).

Inspection of the plots depicting expression of memory-enriched genes revealed that the degree of skewing by both *P* value and fold change appeared to be greater in the Lt-HSC versus LCP than the Lt-HSC versus St-HSC comparison. Indeed, those transcripts whose expression values correlated best with the biggest difference in self-renewal capacity, i.e., those most enriched in Lt-HSC, vis-à-vis LCP, were also enriched in Lt-HSC relative to St-HSC and in St-HSC relative to LCP (Fig. 1*D* in red). Furthermore, the memory-enriched transcripts appeared to be progressively lost with differentiation. Their loss correlated with progressive loss of self-

renewal function in all three progenitor populations. In addition, because St-HSC retain full differentiation potential, the enriched transcripts are unlikely to represent genes solely involved in lymphocyte biology or fate commitment.

We also looked at those transcripts selectively down-regulated in memory CD8$^+$ T cells relative to both naïve and effector cells (Fig. 1 *E* and *F*). One hundred two transcripts were identified with expression levels selectively down-regulated in memory CD8$^+$ cells relative to both naïve and effector cells. Again, we observed a strong correlation between expression trends in memory CD8$^+$ T cells and Lt-HSC. The transcripts that were absent or down-regulated in memory CD8$^+$ cells were also relatively depleted in Lt-HSC (Fig. 1*G*). The transcripts most down-regulated in Lt-HSC versus LCP also appeared to be progressively up-regulated as cells differentiated into St-HSC and LCP (Fig. 1*H* in red). Together with the data presented in Fig. 1 *A–D*, this result demonstrated that the majority of T cell memory enriched transcripts are also enriched in Lt-HSC, whereas the majority of T cell memory depleted transcripts are depleted in Lt-HSC.

Next we addressed whether a similar correlation could be observed between Lt-HSC and the other self-renewing mature lymphocyte population, memory B cells. Thus, we turned to an independently generated collection of Affymetrix GeneChip data on memory, naïve, germinal center, and plasma B cells. To generate antigen-specific B cells *in vivo*, we immunized mice with the T-dependent immunogen NP-CGG, and antigen-specific cells were harvested at various time-points (detailed in *Supporting Methods*, which are published as supporting information on the PNAS web site). These cells were functionally verified by transfer into RAG$^{-/-}$ hosts and measuring T-dependent antibody production (Fig. 6, which is published as supporting information on the PNAS web site).

In a fashion parallel to the above analysis, we compared the gene-expression profiles of memory and naïve B cells (Fig. 2*A*). The transcripts whose levels were most increased in memory cells were selected and subsequently plotted versus germinal center cells (Fig. 2*B*) and then plasma cells (Fig. 2*C*). Thereby, a set of 272 transcripts enriched in memory B cells relative to both naïve, germinal center, and plasma B cells were delineated. Analysis of these enriched transcripts within the various stem-cell populations revealed a preferential representation in Lt-HSC-enriched genes compared with either LCP or St-HSC, both in numbers (71% and 79%, respectively) and *P* values (Fig. 2*D*). Just as was observed in the T cell analysis, the degree of skewing of the memory B cell-enriched transcripts correlated inversely with the progressive loss of self-renewal capacity (Fig. 2*E* in red). Clearly then, a large fraction of those transcripts augmented in memory B cells relative to naïve, germinal center, and plasma cells were also selectively enriched in Lt-HSC.

A set of 481 transcripts selectively down-regulated in memory B cells relative to the other B cell populations was also delineated (Fig. 2 *F–H*). Again, a strong correlation between expression in memory B cells and Lt-HSC was observed, as transcripts depleted in memory B cells were diminished in Lt-HSC relative to both LCP and St-HSC (Fig. 2*I*). Those transcripts most down-regulated in Lt-HSC versus LCP also appeared to be progressively up-regulated as cells differentiated into St-HSC and LCP (Fig. 2*J* in red). These observations demonstrate that the majority of B cell-memory-enriched genes were augmented in Lt-HSC, whereas the majority of B cell-memory-depleted transcripts were depleted in Lt-HSC. This separate B cell data set provides an important independent confirmation of the T cell data comparisons.

Those transcripts whose expression was up-regulated in both memory T and memory B cells relative to their non-self-renewing counterparts were tabulated (Fig. 3*A*). Virtually all (92% and 85%) of these shared transcripts were enriched in Lt-HSC (Fig. 3*B*). Similar results were obtained when those genes down-regulated in both memory populations were compared, with 88% and 84% also

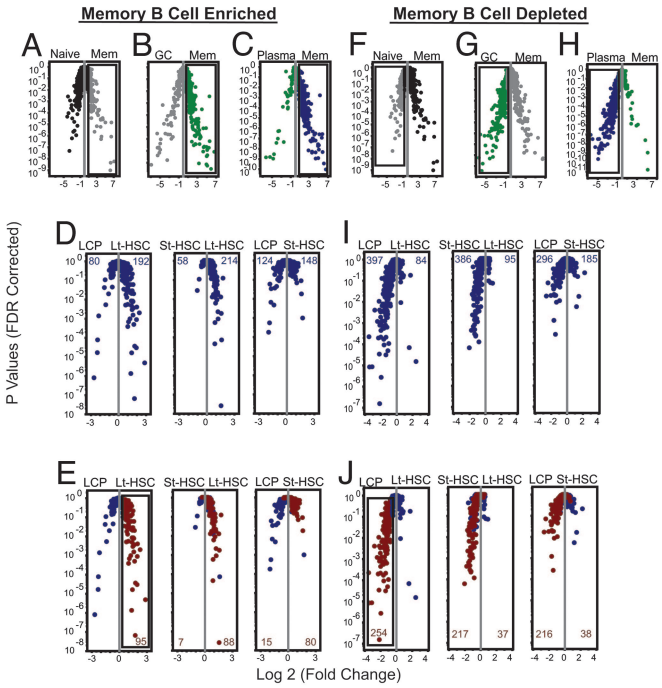

Fig. 2. Memory B cell-enriched transcripts are also enriched in Lt-HSC, whereas memory B cell-depleted transcripts are also depleted in Lt-HSC. The 272 transcripts whose expression was relatively enriched (fold change $\geq$1.4) in memory B cells relative to naïve (*A*), germinal center (*B*), and plasma B cells (*C*) are plotted for their relative expression in Lt-HSC vs. LCP, Lt-HSC vs. St-HSC, and St-HSC vs. LCP (*D*). The number of transcripts whose $\log_2$ (fold change) is greater or less than zero is shown at the top of the plot. Those transcripts whose expression in Lt-HSC vs. LCP was >1.4 were then highlighted in red (*E*), and their number greater or less than zero is shown at the bottom of the plot. The 481 transcipts whose expression was relatively depleted (≤ -1.4) in memory B cells relative to naïve (*F*), germinal center (*G*), and plasma B cells (*H*) are plotted for their relative expression in Lt-HSC vs. LCP, Lt-HSC vs. St-HSC, and St-HSC vs. LCP (*I*). The number of transcripts whose $\log_2$ (fold change) is greater or less than zero is shown at the top of the plot. Those transcripts whose expression in Lt-HSC vs. LCP was <-1.4 were then highlighted in red (*J*), and their number greater or less than zero is shown at the bottom of the plot.

down-regulated in Lt-HSC (Fig. 3*C*). We propose that these transcripts underlie the most restrictive transcriptional definition of immune memory, and virtually all of them were coordinately regulated in Lt-HSC. The transcripts expressed in concert among memory T, memory B, and Lt-HSC likely represent a transcriptional profile of self-renewal in these diverse hematolymphoid cells.

Q-PCR provided confirmation of the expression of several of those transcripts coordinately enriched in memory T cells and Lt-HSC. Lt-HSC were purified as Lin$^{-/lo}$, Sca$^+$, c-kit$^+$, CD34$^-$, and Flt3$^-$ cells (19–21). Likewise, St-HSC were purified as Lin$^{-/lo}$, Sca$^+$, c-kit$^+$, CD34$^+$, and Flt3$^-$ cells and LCP as Lin$^{-/lo}$, Sca$^+$, c-kit$^+$, CD34$^+$, and Flt3$^+$ cells. Although this LCP population differs from the one used in the GeneChip analysis in its Sca expression, both populations lack self-renewing capacity and show a limited differentiation capacity (14, 19–22). In nearly every individual comparison, the Q-PCR data confirmed what was observed in the microarray analysis (Fig. 7, which is published as supporting information on the PNAS web site). Indeed, the differences observed by Q-PCR were often much greater than those estimated from the chip data. Two noticeable exceptions were the cytokine receptors IL-18R and IL-7R. IL-18R did not appear to be enriched in either HSC population, whereas IL-7R transcript levels were high in Lt-HSC, low in St-HSC, and highest in LCP.

To observe how these shared transcripts partitioned in other stem cell comparisons, we considered our T cell data in con-

A *Transcripts Up-regulated in Tmem and Bmem*

Symbol	Name	Unigene	Positive Fold Change		
			Lt-HSC vs. LCP	Lt-HSC vs. St-HSC	St-HSC vs. LCP
Plekha5	Pleckstrin homology domain containing, family A	Mm.80531	2.60	1.88	1.38
B930046C15R	RIKEN cDNA B930046C15 gene	Mm.204	2.30	1.65	1.39
Mapk12	Mitogen-activated protein kinase 12	Mm.38343	2.07	1.55	1.34
BC042396	cDNA sequence BC042396	Mm.24944	2.03	1.39	1.45
Il18r1	Interleukin 18 receptor 1	Mm.4773	1.96	1.29	1.53
Il7r	Interleukin 7 receptor	Mm.389	1.73	1.87	0.93
Apoe	Apolipoprotein E	Mm.138866	1.54	1.19	1.30
Sema4f	Semaphorin 4f	Mm.42035	1.26	1.12	1.13
Pou6f1	POU domain, class 6, transcription factor 1	Mm.28825	1.22	1.07	1.15
Prkcz	Protein kinase C, zeta	Mm.28561	1.10	1.10	1.00
Ebi2	Epstein-Barr virus induced gene 2	Mm.3007	1.09	1.09	1.01
na	Similar to zinc finger protein 187	Mm.24023	1.08	1.23	0.88
Usp48	Ubiquitin specific protease 48	Mm.21824	0.90	1.04	0.86
Fcgr2b	Fc receptor, IgG, low affinity IIb	Mm.10809	0.23	0.73	0.32

Transcripts Down-regulated in Tmem and Bmem

Symbol	Name	Unigene	Negative Fold Change		
			Lt-HSC vs. LCP	Lt-HSC vs. St-HSC	St-HSC vs. LCP
2310008M10R	RIKEN cDNA 2310008M10 gene	Mm.9811	3.48	1.16	3.01
AI317230	EST AI317230	Mm.151329	2.97	1.34	2.20
Coq7	Demethyl-Q 7	Mm.20634	2.62	1.55	1.69
Ywhah	Eta polypeptide	Mm.3308	2.57	1.52	1.68
Myb	Myeloblastosis oncogene	Mm.1202	2.39	0.75	3.20
Eif1ay	Eukaryotic translation initiation factor 1A, Y-linked	Mm.65264	2.23	1.40	1.60
1110004P21Ri	RIKEN cDNA 1110004P21 gene	Mm.43213	2.23	0.78	2.87
3110003A17Ri	RIKEN cDNA 3110003A17 gene	Mm.28149	2.13	1.18	1.80
Cbx3	Chromobox homolog 3 (Drosophila HP1 gamma)	Mm.28148	1.83	1.53	1.19
Sfrs9	Splicing factor, arginine/serine rich 9	Mm.29426	1.66	1.23	1.35
Snrpe	Small nuclear ribonucleoprotein E	Mm.27669	1.63	1.21	1.35
Hnrpa2b1	Heterogeneous nuclear ribonucleoprotein A2/B1	Mm.16767	1.43	1.58	0.91
Gnb1	Guanine nucleotide binding protein, beta 1	Mm.2344	1.37	1.03	1.33
Erh	Enhancer of rudimentary homolog (Drosophila)	Mm.21952	1.35	2.11	0.64
Hnrpa3	Heterogeneous nuclear ribonucleoprotein A3	Mm.219668	1.02	0.99	1.03
2310061N23Ri	RIKEN cDNA 2310061N23 gene	Mm.46382	0.87	1.10	0.79
Hrmt1l2	Heterogeneous nuclear ribonucleoproteins	Mm.27545	0.58	2.85	0.21

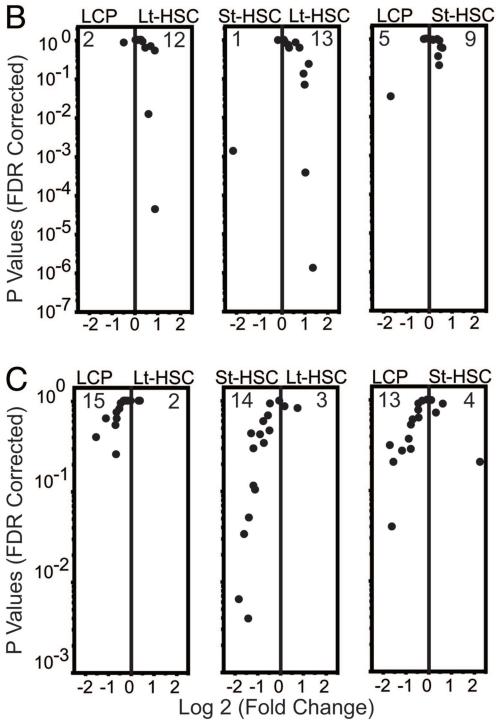

Fig. 3. Coordinately regulated transcripts in both memory CD8$^+$ T cells and memory B cells are also coordinately regulated in Lt-HSC. (*A*) Those transcripts whose expression in coordinately regulated in memory B cells and memory T cells are listed. (*B*) Enriched transcripts are plotted for their relative expression in Lt-HSC vs. LCP, Lt-HSC vs. St-HSC, and St-HSC vs. LCP. (*C*) Depleted transcripts are plotted for their relative expression in Lt-HSC vs. LCP, Lt-HSC vs. St-HSC, and St-HSC vs. LCP. The number of transcripts whose $\log_2$ (fold change) is greater or less than zero is shown at the top of the plots.

junction with other array comparisons of stem cell populations. Although a number of additional hematopoietic stem cell global gene expression studies have been performed (23–25), previous work has shown that cross-platform comparisons do not correlate well (26). Therefore, we focused our analyses on studies performed with Affymetrix-based experiments and used the complete data sets published by Ivanova *et al.* (14), Ramalho-Santos *et al.* (27), and Akashi *et al.* (28). We are aware of only one other recently published Affymetrix data set comparing HSC populations (29), but we were unable to obtain the original files in time for inclusion in this publication. Collectively, the three independent data sets analyzed included adult and fetal hematopoietic stem cells as well as embryonic and neural stem cells. The vast majority of those genes whose expression we identified as coenriched in memory T cells and memory B cells (vis-à-vis their designated counterparts) were also augmented in the adult hematopoietic stem cell populations of Ramalho-Santos *et al.* (27) and Akashi *et al.* (ref. 28; Fig. 4). Furthermore, most of the shared transcripts were also increased in fetal-liver hematopoietic cell precursors of Ivanova *et al.* (14). These findings provide independent confirmation of our results, suggesting that this common "self-renewal" molecular signature might be a general feature of hematopoietic stem cell populations. However, there was not a consistent enrichment of the shared transcripts in either neural stem cell or ES cell populations, arguing that this particular molecular signature may be restricted to the self-renewing cells of the hematopoietic system.

Discussion

We sought to provide biological evidence for or against the hypothesis that memory T and memory B cells have reacquired the expression of molecules characteristic of long-term stem cells, coincident with their ability to self-renew. The data presented demonstrate that for both memory T and B cells, a significant subset of their transcripts was also found in Lt-HSC. Indeed, virtually all of those selected transcripts whose expression was most closely coordinated in B and T memory cells were similarly regulated in Lt-HSC. These observations provide evidence supporting the hypothesis that the self-renewal pathways used in memory T and B cells are related to those of hematopoietic stem cells.

Although nearly all of the transcripts shared between memory B and T cells were also found in Lt-HSC, there were many more transcripts shared between only one memory population and Lt-HSC (Tables 1–3, which are published as supporting information on the PNAS web site). Because hematopoietic stem cells are absolutely critical for the survival of the organism, it is likely that they rely on redundant pathways involved in self-renewal, only some of which are used in a given memory lymphocyte population. One explanation for the limited overlap observed between memory T and memory B cells is that memory T cells have reactivated different self-renewal pathways than memory B cells. Support for this explanation can be found by looking at those transcripts shared between Lt-HSC and only one memory population. Bmi-1 is conserved between memory B cells and Lt-HSC but is not enriched in memory T cells. Bmi-1 is a polycomb family member involved in self-renewal of hematopoietic stem cells (30), leukemia cells (31), and central and peripheral nervous system cells (32). Conversely, memory T cells and Lt-HSC, but not memory B cells, have up-regulated Iex-1 and Spi-2A. These transcripts are known regulators of apoptosis that function in memory CD8$^+$ T cell survival (33–35). These examples support the hypothesis that a given memory cell lineage may have reactivated only a subset of the redundant pathways expressed in HSC.

Still other transcripts shared between one memory population and Lt-HSC have functions consistent with their potentially playing a role in self-renewal. Memory T cells and Lt-HSC share expression of TNF receptor II (p74R) and TNF receptor-associated factor 1 (Traf-1). These proteins associate within the cell, resulting in a

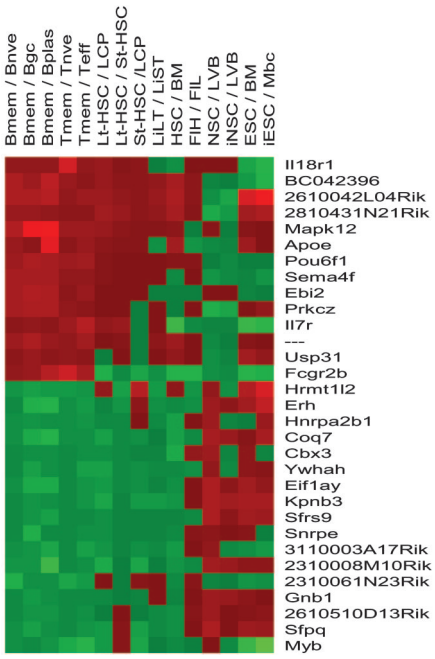

Fig. 4. Coordinately regulated transcripts in memory B and memory T cells are coordinately regulated in several different data sets of hematopoietic stem cells but are not coordinately regulated in neural or ES cells. Transcripts listed in Fig. 5 are on the *y* axis, and their expression is plotted as a heat map in each of the comparisons listed on the *x* axis. Up-regulated transcripts are shown in red, and down-regulated transcripts are shown in green. LiLt and LiSt represent Lt-HSC and St-HSc in the published data of Akashi *et al.* (28) HSC and BM represent the Lt-HSC and mature bone marrow of Ramalho-Santos *et al.* (27). FlH and FIL represent the fetal liver HSC and lineage committed progenitors of Ivanova *et al.* (14). NSC and LVB represent the neural stem cells and lateral ventrical of the brain of Ramalho-Santos *et al.* (27) iNSC represent the neural stem cells of Ivanova *et al.* (14). ESC and BM represent the ES cells and bone marrow of Ramalho-Santos *et al.* (27) iESC and Mbc represent the ES cells and mature bone marrow of Ivanova *et al.* (14). All of the data sets were pooled together for rma analysis as described in *Methods*.

signal that inhibits apoptosis (36, 37). Also present on the shared memory T cell list were several members of the RAS/mitogen-activated protein kinase pathway, known to be involved in decisions by stem cells to undergo proliferation, apoptosis, and differentiation (38–40). Likewise, the memory B cells and Lt-HSC share expression of several classes of transcripts that likely function in self-renewal. Tcf4 and Tcf12 are potentially downstream of β-catenin signaling, itself known to play a role in self-renewal in several stem cell systems (41). Finally, Mef2a and Mef2d are members of a class of transcription factors known to help translate calcium signals in neurons into long-term survival (42, 43). Taken together, the presence of these particular transcripts supports the general hypothesis that memory cells have selectively reactivated different self-renewal molecular pathways found in Lt-HSC.

Even though our data point to different self-renewal pathways being reactivated in either memory B or T cells, we observed several transcripts whose expression was shared between Lt-HSC and both memory populations. Of these jointly shared transcripts, only IL-7R has been shown to play a role in memory T cell self-renewal (5). Although IL-7R clearly plays a role in B cell progenitor differentiation, its role in memory B cell function is unknown. IL-7R is likely to be functionally required for memory B cell self-renewal. However, it is unlikely to function at the level of stem cells, because the IL-7R protein is not expressed on Lt-HSC cell surfaces. A more straightforward explanation for its Lt-HSC expression it that IL-7R gene transcription lies downstream of a common self-renewal pathway. Alternatively, there are several shared transcripts that are more likely to play a functional role in self-renewal. In particular, the signaling molecules mitogen-activated protein kinase 12 and PKC-ζ and the transcription factor Pou6f1 represent potentially convergent nodes in the network of self-renewal pathways.

Identification of these transcripts lends significant impetus for further testing of their functional relevance to hematopoietic and memory cell self-renewal. In particular, our data suggest that the polycomb complex that includes Bmi-1 is likely to function in memory B cell self-renewal in addition to its already reported role in hematopoietic stem cells. Given the role of polycomb genes in the maintenance of cellular memory of chromatin modification and transcriptional repression, these molecules are particularly intriguing candidates for functioning in immunologic memory. Further, it is worth considering the possibility that the separate pathways identified in our analysis might functionally converge within the cell. For instance, Bmi-1 itself has recently been shown to associate with and be phosphorylated by 3pK (mitogen-activated protein kinase AP kinase 3), which lies downstream of several mitogen-activated protein kinase pathways (44).

There has been a great deal of debate concerning the validity and reproducibility of defining a general molecular signature of stem cells (14, 27, 45, 46). Although "stemness" certainly requires an aspect of self-renewal, there are many additional functions and/or states that might be shared by the broad range of stem cells examined in the previous studies. Furthermore, it is not difficult to imagine arriving at similar phenotypes via divergent pathways, particularly within different lineages. Because memory T and memory B cells are descended from long-term and short-term HSC, we suggest that the focused comparisons presented herein provide unique insights into self-renewal within the hematopoietic system. Indeed, those genes shared between both memory populations were coordinately regulated in all three of the published HSC Affymetrix data sets we analyzed in Fig. 6. However, when the T cell data were considered in conjunction with previously published ES cell and neural stem cell data sets (14, 27), there was not a consistent enrichment of the shared transcripts in either of these two. This finding suggests that the molecular signature we defined may be restricted to the self-renewing cells of the hematopoietic system, a finding consistent with the published work of others showing conservation within, but not across, lineages (45, 46).

Our results have important implications beyond the identification of a self-renewal signature. For example, these shared transcripts are excellent candidates for those reactivated in the self-renewal program of leukemic stem cells (47–51). Indeed, an increase in the expression of the polycomb complex component Bmi-1 has been implicated in leukemogenesis (31, 52). Second, given the recent reports that memory $CD8^+$ T cell self-renewal is preferentially localized to the bone marrow (53), it is an intriguing possibility that memory T cells and hematopoietic stem cells may have partially overlapping niches within the marrow that support their self-renewal. Finally, the data provides a glimpse of the shared biochemical mechanisms with which hematopoietic cells undergo self-renewal.

Methods

T Cell Purification, RNA Processing, and Amplification. T cells were sorted, and RNA was purified, amplified, and hybridized as described in ref. 15. Details for purification are given in *Supporting Methods*. Replicates included naïve (four), effector (three), and memory (five) populations.

B Lineage Cell Purification, RNA Processing, and Amplification. The purification strategy, ELISPOT assays, RNA processing method, and hybridization strategy are all detailed in *Supporting Methods*. Replicates included naïve (three), germinal center (three), plasma (four), and memory (four) populations.

Hematopoietic Stem Cell Purification for Q-PCR Confirmation. Lt-HSC, St-HSC, and LCP purification was performed by following the protocol described by Yang *et al.* (20). Details of purification strategy and Q-PCR methods are described in *Supporting Methods*.

Statistical Methods. Affymetrix image files (.cel) of the MGU74vA2-A chips from the stem cell and T cell data sets were collectively analyzed by using the AFFYLMGUI package developed by the open-source collaborative www.bioconductor.org (54). Data were background corrected, probe-level normalized and summarized by using the rma method (55). The rma method uses an improved algorithm for probe-level background correction, normalization, and summary that dramatically reduces observed statistical noise both in published control data sets (55–58) and among replicates within our own data (C.J.L., A.W.G., C.B., and D.M., unpublished results). Differential expression and false discovery rate-corrected *P* values were determined by using the LIMMA method (59). Affymetrix image files (.cel) of the 430.V2 chips from the from the B cell data sets were collectively analyzed by using the AFFYLMGUI package as described above. These data sets were then linked at the probe level with the B and T cell data sets by using the published best match correlation files from Affymetrix.

We thank Amy Wagers for her helpful advice and suggestions; Natalia Ivanova, Ihor Lemischka, Miguel Ramalho-Santos, Douglas A. Melton, and Liheng Li for their help with data acquisition; Vanessa Tran, Quynh-Mai Pham, Vera Bruklich, and Tatyana Lipatova for assistance with mice; and Joyce LaVecchio and Grigoriy Losyev for cell sorting. This work was supported by Joslin's Diabetes and Endocrinology Research Center Cores Grant 2 P30 DK36836-17, National Institutes of Health (NIH) Grant AI51530-01, and the W. T. Young Chair for Diabetes Research (to C.B. and D.M.) and NIH Grants R01HL058770, R01AI047457, and P01DK053074 (to I.L.W.). A.W.G. was supported by an Irvington Institute–Juvenile Diabetes Research Fellowship and Joslin's National Institutes of Health Training Grant 5 T32 DK007260-24. D.B. was funded by a Cancer Research Institute Fellowship. C.J.L. was funded by Brigham and Women's Pathology Department Training Grant T32 HL0762–18 and NIH Grant K08 AI063386.

1. Fearon, D. T., Manders, P. & Wagner, S. D. (2001) *Science* **293,** 248–250.
2. Kaech, S. M., Wherry, E. J. & Ahmed, R. (2002) *Nat. Rev. Immunol.* **2,** 251–262.
3. Lanzavecchia, A. & Sallusto, F. (2002) *Nat. Rev. Immunol.* **2,** 982–987.
4. Kondo, M., Wagers, A. J., Manz, M. G., Prohaska, S. S., Scherer, D. C., Beilhack, G. F., Shizuru, J. A. & Weissman, I. L. (2002) *Annu. Rev. Immunol.*, 759–806.
5. Schluns, K. S. & Lefrancois, L. (2003) *Nat. Rev. Immunol.* **3,** 269–279.
6. Woodland, R. T. & Schmidt, M. R. (2005) *Semin. Immunol.* **17,** 209–217.
7. Macallan, D. C., Asquith, B., Irvine, A. J., Wallace, D. L., Worth, A., Ghattas, H., Zhang, Y., Griffin, G. E., Tough, D. F. & Beverley, P. C. (2003) *Eur. J. Immunol.* **33,** 2316–2326.
8. Sprent, J. (2003) *Microbes Infect.* **5,** 227–231.
9. Forster, I., Vieira, P. & Rajewsky, K. (1989) *Int. Immunol.* **1,** 321–331.
10. Schittek, B. & Rajewsky, K. (1990) *Nature* **346,** 749–751.
11. Smith, K. G., Weiss, U., Rajewsky, K., Nossal, G. J. & Tarlinton, D. M. (1994) *Immunity* **1,** 803–813.
12. Slifka, M. K., Antia, R., Whitmire, J. K. & Ahmed, R. (1998) *Immunity* **8,** 363–372.
13. McHeyzer-Williams, L. J., Cool, M. & McHeyzer-Williams, M. G. (2000) *J. Exp. Med.* **191,** 1149–1166.
14. Ivanova, N. B., Dimos, J. T., Schaniel, C., Hackney, J. A., Moore, K. A. & Lemischka, I. R. (2002) *Science* **298,** 601–604.
15. Goldrath, A. W., Luckey, C. J., Park, R., Benoist, C. & Mathis, D. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 16885–16890.
16. Homann, D., Teyton, L. & Oldstone, M. B. (2001) *Nat. Med.* **7,** 913–919.
17. Wherry, E. J. & Ahmed, R. (2004) *J. Virol.* **78,** 5535–5545.
18. Kaech, S. M., Hemby, S., Kersh, E. & Ahmed, R. (2002) *Cell* **111,** 837–851.
19. Adolfsson, J., Borge, O. J., Bryder, D., Theilgaard-Monch, K., Astrand-Grundstrom, I., Sitnicka, E., Sasaki, Y. & Jacobsen, S. E. (2001) *Immunity* **15,** 659–669.
20. Yang, L., Bryder, D., Adolfsson, J., Nygren, J., Mansson, R., Sigvardsson, M. & Jacobsen, S. E. (2005) *Blood* **105,** 2717–2723.
21. Adolfsson, J., Mansson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C. T., Bryder, D., Yang, L., Borge, O. J., Thoren, L. A., *et al.* (2005) *Cell* **121,** 295–306.
22. Christensen, J. L. & Weissman, I. L. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 14541–14546.
23. Forsberg, E. C., Prohaska, S. S., Katzman, S., Heffner, G. C., Stuart, J. M. & Weissman, I. L. (2005) *PLoS Genet.* **1,** e28.
24. Terskikh, A. V., Miyamoto, T., Chang, C., Diatchenko, L. & Weissman, I. L. (2003) *Blood* **102,** 94–101.
25. Park, I. K., He, Y., Lin, F., Laerum, O. D., Tian, Q., Bumgarner, R., Klug, C. A., Li, K., Kuhr, C., Doyle, M. J., Xie, T., *et al.* (2002) *Blood* **99,** 488–498.
26. Park, P. J., Cao, Y. A., Lee, S. Y., Kim, J. W., Chang, M. S., Hart, R. & Choi, S. (2004) *J. Biotechnol.* **112,** 225–245.
27. Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R. C. & Melton, D. A. (2002) *Science* **298,** 597–600.
28. Akashi, K., He, X., Chen, J., Iwasaki, H., Niu, C., Steenhard, B., Zhang, J., Haug, J. & Li, L. (2003) *Blood* **101,** 383–389.
29. Zhong, J. F., Zhao, Y., Sutton, S., Su, A., Zhan, Y., Zhu, L., Yan, C., Gallaher, T., Johnston, P. B., Anderson, W. F. & Cooke, M. P. (2005) *Proc. Natl. Acad. Sci. USA* **102,** 2448–2453.
30. Park, I. K., Qian, D., Kiel, M., Becker, M. W., Pihalja, M., Weissman, I. L., Morrison, S. J. & Clarke, M. F. (2003) *Nature* **423,** 302–305.
31. Lessard, J. & Sauvageau, G. (2003) *Nature* **423,** 255–260.
32. Molofsky, A. V., Pardal, R., Iwashita, T., Park, I. K., Clarke, M. F. & Morrison, S. J. (2003) *Nature* **425,** 962–967.
33. Liu, N., Phillips, T., Zhang, M., Wang, Y., Opferman, J. T., Shah, R. & Ashton-Rickardt, P. G. (2004) *Nat. Immunol.* **5,** 919–926.
34. Liu, N., Wang, Y. & Ashton-Rickardt, P. G. (2004) *FEBS Lett.* **569,** 49–53.
35. Zhang, Y., Schlossman, S. F., Edwards, R. A., Ou, C.-N., Gu, J. & Wu, M. X. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 878–883.
36. Zapata, J. M. & Reed, J. C. (2002) *Sci. STKE* **2002** (133), PE27.
37. Baud, V. & Karin, M. (2001) *Trends Cell Biol.* **11,** 372–377.
38. Dorrell, C., Takenaka, K., Minden, M. D., Hawley, R. G. & Dick, J. E. (2004) *Mol. Cell. Biol.* **24,** 6993–7002.
39. Qi, X., Li, T.-G., Hao, J., Hu, J., Wang, J., Simmons, H., Miura, S., Mishina, Y. & Zhao, G.-Q. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 6027–6032.
40. Burdon, T., Smith, A. & Savatier, P. (2002) *Trends Cell. Biol.* **12,** 432–438.
41. Reya, T. & Clevers, H. (2005) *Nature* **434,** 843–850.
42. Mao, Z., Bonni, A., Xia, F., Nadal-Vicens, M. & Greenberg, M. E. (1999) *Science* **286,** 785–790.
43. Gaudilliere, B., Shi, Y. & Bonni, A. (2002) *J. Biol. Chem.* **277,** 46442–46446.
44. Voncken, J. W., Niessen, H., Neufeld, B., Rennefahrt, U., Dahlmans, V., Kubben, N., Holzer, B., Ludwig, S. & Rapp, U. R. (2005) *J. Biol. Chem.* **280,** 5178–5187.
45. Evsikov, A. V. & Solter, D. (2003) *Science* **302,** 393; author reply 393.
46. Fortunel, N. O., Otu, H. H., Ng, H. H., Chen, J., Mu, X., Chevassut, T., Li, X., Joseph, M., Bailey, C., Hatzfeld, J. A., *et al.* (2003) *Science* **302,** 393; author reply 393.
47. Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M. A. & Dick, J. E. (1994) *Nature* **367,** 645–648.
48. Sirard, C., Lapidot, T., Vormoor, J., Cashman, J. D., Doedens, M., Murdoch, B., Jamal, N., Messner, H., Addey, L., Minden, M., *et al.* (1996) *Blood* **87,** 1539–1548.
49. Jamieson, C. H., Ailles, L. E., Dylla, S. J., Muijtjens, M., Jones, C., Zehnder, J. L., Gotlib, J., Li, K., Manz, M. G., Keating, A., *et al.* (2004) *N. Engl. J. Med.* **351,** 657–667.
50. Passegue, E., Jamieson, C. H. M., Ailles, L. E. & Weissman, I. L. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 11842–11849.
51. Warner, J. K., Wang, J. C., Hope, K. J., Jin, L. & Dick, J. E. (2004) *Oncogene* **23,** 7164–7177.
52. Raaphorst, F. M. (2003) *Trends Immunol.* **24,** 522–524.
53. Di Rosa, F. & Pabst, R. (2005) *Trends Immunol.* **26,** 360–366.
54. Wettenhall, J. M. & Smyth, G. K. (2004) *Bioinformatics* **20,** 3705–3706.
55. Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U. & Speed, T. P. (2003) *Biostatistics* **4,** 249–264.
56. Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B. & Speed, T. P. (2003) *Nucleic Acids Res.* **31,** e15.
57. Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. (2003) *Bioinformatics* **19,** 185–193.
58. Wu, Z. & Irizarry, R. A. (2004) *Nat. Biotechnol.* **22,** 656–658; author reply 658.
59. Smyth, G. K. (2004) *Statistical Applications in Genetics and Molecular Biology*, Vol. 3, Issue 1, article 3. Available at www.bepress.com/sagmb/vol3/iss1/art3. Accessed January 23, 2006.