Generation of human memory stem T cells upon haploidentical T-replete hematopoietic stem cell transplantation

Nicoletta Cieri1,2,3, Giacomo Oliveira1,3, Raffaella Greco2,3, Mattia Forcato4, Cristian Taccioli4, Beatrice Cianciotti1, Veronica Valtolina1, Maddalena Noviello1, Luca Vago2,5, Attilio Bondanza1,2,3, Francesca Lunghi2, Sarah Marktel2, Laura Bellio6, Claudio Bordignon3,7, Silvio Bicciato4, Jacopo Peccatori2, Fabio Ciceri2,5,6† and Chiara Bonini1*†

Affiliations
1 Experimental Hematology Unit, Division of Immunology, Transplantation and Infectious Diseases, Program In Immunology and Bio-immunotherapy of Cancer (PIBIC), San Raffaele Scientific Institute, Milan, Italy.
2 Hematology and Bone Marrow Transplantation Unit, Department of Oncology, Division of Regenerative Medicine, Stem Cells and Gene Therapy, San Raffaele Scientific Institute, Milan, Italy.
3 University Vita-Salute San Raffaele, Milan, Italy.
4 Center for Genome Research, Department of Life Sciences, University of Modena and Reggio Emilia, Modena, Italy.
5 Unit of Molecular and Functional Immunogenetics, Division of Regenerative Medicine, Stem Cells and Gene Therapy, San Raffaele Scientific Institute, Milan, Italy.
6 Immunohematology and Transfusion Medicine Unit, Division of Regenerative Medicine, Stem Cells and Gene Therapy, San Raffaele Scientific Institute, Milan, Italy.
7 Molmed S.p.A., Milan, Italy

*To whom correspondence should be addressed: Dr Chiara Bonini, Experimental Hematology Unit, San Raffaele Scientific Institute, via Olgettina 60, Milano, Italy. e-mail: bonini.chiara@hsr.it.
† These authors contributed equally to the work.

Scientific category: Immunobiology (Secondary scientific category: Transplantation)
Running Title: T_{SCM} generation upon haploidentical HSCT
Abstract

Memory stem T cells (T\textsubscript{SCM}) have been proposed as key determinants of immunological memory. However, their exact contribution to a mounting immune response as well as the mechanisms and timing of their \textit{in vivo} generation are poorly understood. We longitudinally tracked T\textsubscript{SCM} dynamics in patients undergoing haploidentical hematopoietic stem cell transplantation (HSCT), thereby providing novel hints on the contribution of this subset to post transplant immune reconstitution in humans. We found that donor-derived T\textsubscript{SCM} cells are highly enriched early after HSCT. We showed at the antigen-specific and clonal level that T\textsubscript{SCM} lymphocytes can differentiate directly from naïve precursors infused within the graft and that the extent of T\textsubscript{SCM} generation might correlate with IL-7 serum levels. \textit{In vivo} fate mapping through TCR sequencing allowed defining the \textit{in vivo} differentiation landscapes of human naïve T cells, supporting the notion that progenies of single naïve cells embrace disparate fates \textit{in vivo}, and highlighting T\textsubscript{SCM} as relevant novel players in the diversification of immunological memory following allogeneic HSCT.

Key points

- T\textsubscript{SCM} lymphocytes are preferentially generated from naïve precursors \textit{in vivo} early after haploidentical HSCT;
- T\textsubscript{SCM} cells represent relevant novel players in the diversification of immunological memory following haploidentical HSCT.
**Introduction**

Upon antigen recognition, naive T cells (T$_N$) undergo extensive functional and phenotypic changes that drive their differentiation into effectors, committed to rapidly clear the pathogen, and memory cells, able to survive long term, to ensure recall responses in case of pathogen recurrence\(^1\). The memory compartment is multi-faceted and encompasses multiple T-cell subsets with divergent properties\(^2\). In addition to central memory (T$_{CM}$) and effector memory (T$_{EM}$) cells\(^3\), the spectrum of immunological memory has been recently extended with the identification of memory stem T cells (T$_{SCM}$)\(^4,5\). Human T$_{SCM}$ lymphocytes express CD45RA, CCR7 and CD62L like T$_N$ cells, but differently from T$_N$ and similarly to other memory subsets, are characterized by CD95 expression. Gene expression profiling, corroborated by *in vitro* and *in vivo* experimental results, posits T$_{SCM}$ upstream T$_{CM}$ and T$_{EM}$ in T-cell ontogeny\(^4,5\). However, whether T$_{SCM}$ represent a stable subset, reproducibly generated upon T$_N$ priming, and which are the relationships of T$_{SCM}$ with the other memory subsets are still controversial issues. Furthermore, how human T$_{SCM}$ cells are formed *in vivo*, which are the instructive signals guiding their formation and/or expansion, and when they emerge during the immune response has not been elucidated yet.

The contribution of different memory T-cell subsets (though not T$_{SCM}$) to the generation of effective primary and secondary immune responses has been elegantly studied at single-cell level in murine models\(^6-10\). In humans, however, this pursuit has been limited by ethical and technical constraints. Hematopoietic stem cell transplantation (HSCT) provides a unique setting to longitudinally study the dynamics of discrete T-cell populations upon transfer directly in humans. During allogeneic T-replete HSCT, the recipient receives a myeloablative and lymphodepleting conditioning that abates the host hematopoietic compartment, followed by the
infusion of a graft containing not only hematopoietic stem cells but also mature T cells from an allogeneic donor\textsuperscript{11}. The conditioning regimen induces systemic inflammation, thereby creating a milieu in which infused donor-derived lymphocytes are exposed to a high (allo)antigenic load. Therefore, close longitudinal sampling of peripheral blood (PB) and bone marrow (BM) of patients after HSCT may allow tracking primary and/or secondary T-cell responses. We have previously shown that IL-7 and IL-15 are necessary to instruct the generation of human T\textsubscript{SCM} lymphocytes from naïve precursors \textit{in vitro} and that T\textsubscript{SCM} cells, which represent not more than 2-3\% of circulating T cells in healthy subjects, are enriched early after HSCT, when these homeostatic cytokines are particularly abundant\textsuperscript{5}. Building on these observations, we investigated when and how T\textsubscript{SCM} emerge upon transplant \textit{in vivo}, in the clinically relevant setting of haploidentical HSCT with post transplant cyclophosphamide (PT-Cy)\textsuperscript{12,13}. Understanding the complex T\textsubscript{SCM} dynamics upon allogeneic HSCT may ultimately inform on the role of this novel T-cell subset in post transplant immune reconstitution (IR), and possibly provide novel hints on their physiological role.

**Methods**

\textit{Patients, procedures and biological samples.} Twenty consecutive adult patients with high-risk hematologic malignancies treated with myeloablative haploidentical HSCT in our Hematology Unit were studied. Patients received a treosulfan-based myeloablative conditioning followed by an unmanipulated peripheral blood graft (PBSCs) from an HLA-haploidentical related donor. Postgrafting graft-versus-host-disease (GVHD) prophylaxis consisted of post-transplant cyclophosphamide (PT-Cy, 50 mg/kg/day) on days 3 and 4, followed by mycophenolate mofetil (15 mg/kg 3 times a day) for 30 days, and sirolimus (target level: 8-15 ng/ mL) for 3 months
(supplemental Figure 1A). PB and BM samples of donors, patients and healthy subjects were collected after written informed consent approved from the San Raffaele institutional ethical committee.

**Multiparametric Flow Cytometry.** Mononuclear cells (PBMCs) were isolated from PB and BM by Ficoll-Hypaque gradient separation (Lymphoprep; Fresenius) and immediately used for subsequent analyses. All phenotypic and functional analyses were performed on freshly isolated samples. Eleven-color immunophenotypic analysis was performed using a LSR Fortessa flow cytometer (BD Biosciences). Each acquisition was calibrated using Rainbow Calibration Particles (Spherotech) to correct for day-to-day variation. For intracellular staining, cells were stained with appropriate surface antibodies, washed, and then fixed and permeabilized with FOXP3 Fix/Perm Buffer Set (Biolegend) following manufacturer instructions. Data were processed using FlowJo version 9.7.5 (TreeStar). T-cell subsets were defined according to the expression of CD45RA, CD62L and CD95 as follows: 1) T\(N\): CD45RA\(^+\)CD62L\(^+\)CD95\(^-\); 2) T\(_{SCM}\): CD45RA\(^+\)CD62L\(^+\)CD95\(^-\); 3) T\(_{CM}\): CD45RA\(^-\)CD62L\(^+\); 4) T\(_{EM}\): CD45RA\(^-\)CD62L\(^-\); and 5) T\(_{EMRA}\): CD45RA\(^+\)CD62L\(^-\). Supplemental Material provides detailed information on absolute quantifications, antibodies, fluorochromes, FACS-sorting and gating strategy.

**Aldefluor assay to determine ALDH activity.** Aldefluor kit (Stem Cell Technologies) was used following manufacturer’s instructions, to identify cell populations with high ALDH enzymatic activity. Briefly, PBMCs from leukapheresis or from PB were incubated with Aldefluor substrate for 30 min at 37°C, with and without diethylaminobenzaldehyde (DEAB), and then stained for surface markers. Aldefluor was detected in the green fluorescence channel, and samples treated with DEAB were used as controls to set the gates defining the ALDH\(^+\) region.
**Quantification of serum cytokines.** IL-7 and IL-15 serum concentrations were quantified by Bio-Plex Pro Human Cytokine 4-plex array (Bio-Rad). Samples were analyzed in duplicate. For each studied cytokine, a high-sensitivity standard curve was prepared by serial dilutions of recombinant proteins. Data were analyzed using Bio-Plex Manager version 6.1 (Bio-Rad).

**TCRB CDR3 region sequencing and analysis.** TCRβ chain hypervariable complementarity-determining region 3 (TCRB CDR3) were amplified and sequenced from DNA extracted from T-cell subsets (T_N, T_SCM, T_CM and T_EM/EFF) FACS-purified (purity>95%) from leukapheresis and from PB harvested 30 days post-HSCT from three consecutive patients using the ImmunoSEQ platform at Adaptive Biotechnologies (Seattle, WA, USA)\textsuperscript{14}. TCRB CDR3 region was defined as established by the International ImMunoGeneTics collaboration\textsuperscript{15}. Sequences that did not match CDR3 sequences were removed from the analysis. Rearranged CDR3 sequences were classified as nonproductive if they included insertions or deletions resulting in frameshifts or premature stop codons, and were excluded from subsequent analyses, according to the ImmunoSEQ validated algorithm. The average numbers of total productive nucleotide sequences retrieved were the following: 1) LP samples: T_N: 1,197,527; T_SCM: 1,140,816; T_CM: 2,259,613; T_EM/EFF: 1,416,086; and 2) day 30 samples: T_N: 1,421,741; T_SCM: 1,157,884; T_CM: 1,288,319; T_EM/EFF: 1,788,812. Unique CDR3 sequences were identified and compared using R version 3.0.0. For T-cell subsets harvested 30 days post-HSCT, only CDR3 TCRB sequences whose reads were >10 were considered for all LP-samples and for the qualitative analyses of day 30 samples (to avoid inclusion of reading errors). No filters were applied to day 30-samples for quantitative analyses to ensure the inclusion of all substantial count reductions, indicative of clone contraction upon infusion. TCRB sequence counts and T-cell subsets were clustered using Spearman correlation
and average linkage; heatmaps were drawn using gplots R package. Circos plots were performed as previously described16.

**Statistical analysis.** Statistical analyses were performed with Prism 5 (GraphPad Software). Data are shown as mean ± S.E.M., unless otherwise specified. Data were analyzed with paired t test when comparing the same subset between different time-points or with unpaired t test or standard ANOVA when comparing across two or more different subsets. For all comparisons, two-sided P values were used, and P<0.05 was considered statistically significant. For linear regression analyses, the best-fit values of the slopes were reported along with goodness-of-fit value (r²), P value of the slope (F test) and 95% confidence band. Significance of the slopes was independently verified using R.

**Results**

Donor-derived T\textsubscript{SCM} are selectively enriched early after haploidentical HSCT

To understand the dynamics and fate of human naïve and memory T lymphocytes upon HSCT, donor-derived T-cell subsets were longitudinally tracked within the first month upon infusion into twenty consecutive high-risk hematological patients undergoing haploidentical myeloablative T-replete HSCT with PT-Cy. Patient and donor characteristics are summarized in supplemental Table 1. Of notice, this HSCT platform (supplemental Figure 1A) allows tracking of CD3\textsuperscript{+} T cells even in the first two weeks post-HSCT, a possibility precluded in transplants exploiting anti-thymocyte globulins as in vivo T-cell depleting agent (supplemental Figure 1B). First, to rule out the possible contribution of residual host-derived lymphocytes to circulating cells, we analyzed peripheral T-cell chimerism in 4 patient-donor couples mismatched for HLA-
A*0201. Both in CD4+ and CD8+ compartments, host-derived T cells accounted for less than 30% of cells at day 1 after infusion, decreased to less than 10% at day 5, to disappear by day 15 post-HSCT (supplemental Figure 2). Therefore, donor-derived T cells dominated even at earliest time-points post-HSCT. Upon infusion and administration of PT-Cy, circulating T_N cells gradually contracted and, by day 8 post-HSCT, the vast majority of circulating T lymphocytes were characterized by a memory phenotype (Figure 1A-B), indicating either a preferential sensitivity of T_N to Cy or a massive differentiation of T_N cells into memory/effector cells in the first week after transplant. Strikingly, at day 8 post-HSCT, CD8+ T_{SCM} became the most represented circulating subset, both in terms of frequency and absolute counts (Figure 1C-D). Their frequency at day 8 was significantly higher than that within the infused graft (Figure 1D). CD4+ T-cell subpopulations displayed superimposable dynamics (supplemental Figure 3). In line with T_{SCM} phenotypic characterization previously reported, CD8+ and CD4+ T_{SCM} at day 8 post-HSCT homogeneously expressed CCR7, CXCR4, CD27 and CD28 (supplemental Figure 4). Notably, serum concentration of IL-7 at day 1 appeared to correlate with the number of circulating CD8+ and CD4+ T_{SCM} lymphocytes measured at day 8 (Figure 1E and supplemental Figure 3), but not with T_CM or T_TEM (supplemental Figure 5). At later time-points, T_CM, T_TEM and effectors prevailed, while CD8+ and CD4+ T_N reappeared at day 30 post-HSCT (Figure 1 and supplemental Figure 3). These results prompted us to unveil the origin of those T_{SCM} lymphocytes highly represented at day 8 post-HSCT.

*PT-Cy administration eliminates proliferating T cells but does not prevent T_{SCM} accumulation*

To discern whether the high frequency of T_{SCM} observed at day 8 was due to expansion of T_{SCM} infused within the graft or to their direct *in vivo* generation, we analyzed T-cell subset
proliferation, assessed by Ki-67 staining\textsuperscript{17}. We hypothesized that $T_{SCM}$ expansion should have been preceded by proliferation of either $T_{SCM}$ or $T_N$ lymphocytes, assuming dedifferentiation of $T_{CM}$ or $T_{EM}$ unlikely\textsuperscript{4}. The accumulation of $T_{SCM}$ in this clinical context should also require such proliferating $T_{SCM}$ or $T_N$ lymphocytes to survive Cy, either for intrinsic resistance, or due to temporal uncoupling of proliferation and Cy administration.

At day 3 upon \textit{in vivo} infusion, in the absence of immunosuppressive drugs, all memory CD$8^+$ T-cell subsets robustly proliferated, with $T_{SCM}$ cells displaying significantly higher frequencies of Ki-67$^+$ cells compared to all other subsets (Figure 2A-C). In contrast, only a minority of $T_N$ cells were Ki-67$^+$, in line with the longer timeframe required for priming (Figure 2A-C). CD$4^+$ T-cell subsets displayed superimposable proliferation kinetics, albeit globally to a lower extent than CD$8^+$ (supplemental Figure 6). Cyclophosphamide is a cytotoxic compound active on proliferating cells, used to purge alloreactive T cells \textit{in vivo} and prevent GVHD. T-cell proliferation was abrogated after PT-Cy in all subsets (Figure 2D), suggesting effective killing of cycling lymphocytes. However, T-cell counts did not drop between day 5 and 8 post-HSCT (Figure 2B, grey line, right Y axis), and $T_{SCM}$ cells increased in percentages and counts (Figure 1). These results could be explained either by a selective resistance of cycling $T_{SCM}$ to Cy or by \textit{de novo} generation of $T_{SCM}$. We thus explored whether $T_{SCM}$ cells were resistant to PT-Cy. We failed to record ALDH activity in three patients selected according to sample availability, the major mechanism of Cy inactivation\textsuperscript{18,19}, in total CD$3^+$ T cells neither within the graft nor upon \textit{in vivo} infusion, in contrast with what observed for coexisting CD$34^+$ HSCs, analyzed as internal controls (Figure 2, E and F). Accordingly, we detected high percentages of apoptotic cells within all memory subsets at day 5 post-HSCT (1 day after PT-Cy). Noticeably, at day 8 significantly lower percentages of CD$8^+$ $T_{SCM}$ cells were Annexin V$^+$ compared to $T_{CM}$, $T_{EM}$ and $T_{EMRA}$, while
by day 15 post-HSCT all memory subsets displayed very low levels of apoptosis (Figure 2H). Altogether, these results suggest that non-apoptotic day 8 T_{SCM} cells might stem from \textit{de novo} differentiation from T_{N} lymphocytes.

\textit{CD4^{+} T_{SCM} preferentially retain CD31 expression by day 8 post-HSCT}

To gain further insight on the possible \textit{in vivo} differentiation of infused T_{N} lymphocytes into T_{SCM}, we first exploited the peculiar kinetics of CD31 expression. CD4^{+} T_{N} cells recently exited from thymus can be identified according to the expression of CD31, that is gradually lost upon proliferation\textsuperscript{20}. Importantly, in healthy adults only a minority of circulating CD4^{+} T_{SCM}, T_{CM} and T_{EM} lymphocytes express CD31 (Figure 3A). We reasoned that upon \textit{in vivo} priming of donor-derived CD4^{+} T_{N} cells, CD31 expression could be detected in their progeny if the analyzed timepoint (day 8 post-HSCT) was sufficiently close to the priming to avoid excessive dilution, thereby representing a tool to track T_{N} differentiation by flow cytometry. To verify this hypothesis, purified CD31^{+} CD4^{+} T_{N} cells from three healthy subjects were CFSE-labeled and polyclonally stimulated in the presence of IL-7 and IL-15. Eight days after stimulation, CD31 expression was retrieved on the post-mitotic progeny of stimulated T_{N} cells, which had acquired memory features (supplemental Figure 7), indicating that CD31 downregulation kinetics on RTEs can be exploited to track their fate \textit{in vivo}. Therefore, we analyzed the kinetics of CD31 expression on CD4^{+} T-cell subsets in our cohort of patients (Figure 3C-E). Similarly to healthy controls, within the infused graft a significantly higher fraction of CD4^{+} T_{N} expressed CD31 compared to all memory subsets, including T_{SCM}, and the percentages (and expression levels) of CD31^{+} T_{N}, T_{SCM}, T_{CM} and T_{EM} did not vary till day 3 post-HSCT (Figure 3C-E). Notably, at day 8, when T_{N} cells disappeared from circulation (Figure 3D), CD31^{+} clustered predominantly
within the T_{SCM} compartment, suggesting recent differentiation from T_{N} precursors. By day 30 post-HSCT, expression of CD31 on CD4{sup} T_{SCM} was significantly reduced as compared to day 8, while CD31{sup} T_{N} cells were again detectable, in line with de novo generation, confirmed by sjTREC analysis performed on day 30 cells (supplemental Figure 8). Overall, these data suggest that CD31{sup}CD4{sup} T_{SCM} cells observed at day 8 post-HSCT preferentially derive from CD31{sup}CD4{sup} T_{N} infused within the graft.

Conversion of CD8{sup} T_{N} into T_{SCM} can be traced at the antigen-specific level

To trace the differentiation of antigen-specific CD8{sup} T_{N} into T_{SCM}, we exploited HLA-A{sup}0201 and -A{sup}0101 dextramer complexes loaded with known immunogenic peptides from WT1{sup} or PRAME{sup} in PB or BM of leukemic patients. Tumor-specific T cells are usually rare in healthy subjects and, when detectable, invariably display a naïve phenotype (4). Indeed, in UPN#4 rare HLA-A{sup}0101 restricted WT1-specific T cells retrieved in the donor LP displayed a T_{N} phenotype. By day 30 after HSCT WT1-specific T cells displayed memory features, and were traceable both in PB and at disease site (BM) with T_{SCM} being detectable in the latter (Figure 4A). The same result was obtained in UPN#13, mismatched with the donor for HLA-A{sup}0201. We traced tumor-specific T cells in 3 additional patients receiving a transplant matched for the restriction element analyzed. Of notice, tumor-specific T cells detected in LP displayed a T_{N} phenotype, while after infusion, anti-tumor T cells acquired a memory phenotype in all patients, and comprised T_{SCM} lymphocytes, although their relative contribution varied across different anatomical sites and different patients (Figure 4B). In our series, only one CMV seropositive HLA-B{sup}0702{sup} patient was transplanted from a CMV seronegative donor, allowing to potentially trace a viral-specific primary immune response. Accordingly, pp65-specific T cells in LP
displayed a naïve phenotype. We observed the appearance of pp65-specific CD8+ T\textsubscript{SCM}, in concomitance with viral reactivation (supplemental Figure 9). Overall, these observations support direct conversion of human CD8+ T\textsubscript{N} into T\textsubscript{SCM} cells \textit{in vivo}.

\textit{In vivo fate mapping by TCR clonotyping uncovers the differentiation landscapes of T\textsubscript{N} cells after transplant}

To assess the relevance of \textit{in vivo} differentiation of T\textsubscript{SCM} from T\textsubscript{N} cells upon transplant, we exploited the T-cell receptor (TCR) sequences harbored by individual T lymphocytes as surrogate clonal markers to track T-cell fate upon infusion. CD3+ T\textsubscript{N}, T\textsubscript{SCM}, T\textsubscript{CM} and T\textsubscript{EM/EFF} were FACS-purified from LP and PBMCs harvested 30 days after HSCT in three consecutive patients and subjected to TCRB CDR3 deep sequencing. First, we observed that the TCRB repertoire detected at day 30 in T\textsubscript{N} did not overlap with that of the other T-cell subsets retrieved at day 30 (supplemental Figure 10), suggesting that the majority of day 30 T\textsubscript{N} cells arise from \textit{de novo} generation. For longitudinal evaluation, CDR3 TCRB in frame nucleotide sequences unique to each T-cell subset within LP samples were identified. These specific rearrangements were used to estimate qualitatively and quantitatively the fate after infusion of T cells bearing a specific differentiation phenotype within LP. As expected, sequences harbored by T\textsubscript{N} lymphocytes within LP were retrieved in all memory subsets at day 30, showing that after HSCT T\textsubscript{N} are able to generate the complete spectrum of immunological memory, including T\textsubscript{SCM}, that represented on average 32% of memory T cells (Figure 5A), implying that during T\textsubscript{N} differentiation upon transplant a high number of T\textsubscript{SCM} is generated. Similarly, TCRB sequences originally harbored by LP-T\textsubscript{SCM} were found in all memory subsets at day 30 post-HSCT. Conversely, sequences present in LP-T\textsubscript{CM} were retrieved preferentially in T\textsubscript{CM} and T\textsubscript{EM/EFF}. 
repertoires at day 30 post-HSCT, while only less than 10% were found in the T_{SCM} repertoire. Sequences harbored by LP-T_{EM/EFF} were found mainly within day 30 T_{EM/EFF}, although conversion into T_{CM} was also observed (27%). Again, less than 10% of TCRB sequences identified in LP-T_{EM/EFF} were retrieved in the T_{SCM} compartment at day 30 (Figure 5A). These data are consistent with a preferential progressive framework of differentiation, following the pathway $\text{T}_N \rightarrow \text{T}_{SCM} \rightarrow \text{T}_{CM} \rightarrow \text{T}_{EM/EFF}$ but still allowing for a certain degree of plasticity within each memory subset. Quantitative analysis revealed that only clonotypes originally harbored by LP-T_N were significantly increased in counts at day 30 post-HSCT, regardless of the subsets in which they were identified after transplant (Figure 5B depicting one representative patient and supplemental Figures 11-12 for the remaining 2 patients analyzed). In accordance with flow-cytometric data (Figure 2), sequences unique to LP-T_{SCM}, LP-T_{CM} or LP-T_{EM/EFF} were similar or reduced in counts at day 30 post-HSCT compared to LP, indicating that PT-Cy efficiently dampened their expansion and that possibly only memory lymphocytes that remained unchallenged upon \textit{in vivo} infusion could survive PT-Cy (Figure 5B and supplemental Figures 11-12). T-cell responses are characterized by the clonal burst of the ancestor lymphocyte recognizing its antigen. Given the absence of expansion upon infusion, LP-T_{SCM}, LP-T_{CM} and LP-T_{EM/EFF} clonotypes were not considered informative enough to study memory T-cell dynamics in this transplant setting, and we focused on LP-T_N cells. To gain insight into the differentiation routes adopted by T_N cells upon priming, we evaluated the overlap of the TCRB repertoire from LP-T_N with that of memory subsets retrieved at day 30 post-HSCT. The majority (on average 80%) of clonotypes of LP-T_N lymphocytes retrieved 30 days after transplant displayed a unique differentiation phenotype, with T_{SCM} and T_{CM} being more represented than T_{EM/EFF}. Only 12% of LP-T_N TCRB sequences were found in multiple memory subsets (Figure
5C-D and supplemental Figure 13). By focusing on those LP-TN-derived sequences that were shared among two or more memory subsets at day 30 post-HSCT, we found that approximately 25% of TCRB sequences were shared among day 30 T_{SCM}, T_{CM} and T_{EM/EFF}, although all other possible combinations could be retrieved (Figure 5D-E). The frequency of memory T cell clones originating from individual LP-TN cells was highly variable, and proportional to the level of memory subset diversification retrieved at day 30 post-HSCT (Figure 5F), suggesting that single-cell progenies differentiating into multiple subsets were those characterized by the highest clonal burst. Although the relative contribution of single versus multiple fates of TN cells could not be determined, our in vivo fate mapping suggests that in the HSCT context progenies of single TN cells might embrace disparate fates, with phenotypic and functional heterogeneity achieved at population level. Thus, we propose a revision of the linear differentiation model to take into account the discordance in fate of individual TN cells and the intrinsic plasticity of the memory compartment, leading to a branched differentiation model (Figure 5G).

**Discussion**

While stem cell engraftment and hematological reconstitution are carefully monitored during the first month after HSCT, the shape of T-cell dynamics within this timeframe has remained largely unknown so far. Here, we exploited the PT-Cy platform that differently from other HSCT protocols, mainly relying on the infusion of anti-thymocyte globulins, allows a thorough analysis of circulating cells in this delicate early post-transplant phase to explore the contribution of T_{SCM} cells to post-transplant IR. Indeed, we observed that donor-derived T_{SCM} lymphocytes are highly enriched early after HSCT and we showed, at the polyclonal, antigen-specific and clonal level, that T_{SCM} lymphocytes arising soon after HSCT preferentially derive
from differentiation of T_N infused within the graft, while most memory lymphocytes are purged by PT-Cy. The observation that Cy differentially affects naïve and memory T cells in vivo, due to their different activation kinetics, might provide novel hints on the cellular determinants of alloreactivity after transplant, a matter of intense debate in the field. Indeed, our results, crossed with the high clinical efficacy of PT-Cy in preventing GVHD after HLA-matched and haploidentical HSCT, suggest that alloreactivity in humans largely segregates with infused memory lymphocytes and is thereby likely fueled by cross reactivity. On the other hand, our results imply that the surviving progenies of T_N cells in this specific HSCT setting, acquire a leading role in guiding IR. Given the observed high fraction of T_N that differentiate into T_SCM cells after transplant, it will be important to include CD95 in future studies aiming at deciphering IR kinetics following allogeneic HSCT, to avoid T_N overestimation and to quantify the contribution of T_SCM.

Notably, we showed that IL-7 serum levels measured one day after graft infusion seems to correlate with circulating T_SCM counts at day 8, thereby suggesting that IL-7 might be important for T_SCM generation not only in vitro but also in vivo. The cytokine milieu in which T cells are activated may differently promote effector versus memory differentiation. We could speculate that the increase in circulating IL-7 levels characteristic of the early phase post-HSCT, and associated to the lymphopenic environment, may have favored T_SCM generation in this specific context, as also suggested by murine studies. Thus, it remains to be evaluated whether these results could be confirmed in other relevant human settings, such as longitudinal studies on vaccinations or chronic infections, where IL-7 is not elevated and other possibly confounding variables, such as post-grafting pharmacological immune suppression with mycophenolate mofetil and sirolimus which could influence T-cell metabolism and differentiation, are not
present. Of notice, Roberto et al. report in this issue of Blood that in a similar transplant setting transferred naïve T cells rapidly differentiate into T_{SCM}, which subsequently reconstitute the post-transplant T-cell compartment. Importantly, this work and the one from Roberto et al. rely on two separate models of haploidentical HSCT with major differences in the conditioning regimen used (myeloablative vs. non-myeloablative), stem cell source (peripheral blood vs. bone marrow), post-grafting GVHD prophylaxis (mTOR inhibitors vs. calcineurin inhibitors) and disease background (acute leukemia vs. lymphoma). Regardless of these important differences, both studies ultimately reach a similar model of human T-cell differentiation, with a predominant role played by T_{SCM} lymphocytes, thus suggesting that T_{SCM} generation after transplant in not deeply influenced by any of these variables. It remains to be elucidated whether such an IR profile could be observed also in pediatric patients undergoing haploidentical HSCT with PT-Cy.

Several lines of evidence support the relevance of our findings in other clinical contexts. Approximately 30% of virus-specific memory cells in yellow fever vaccinated subjects have been reported to express CD45RA and CCR7 and to persist for years\textsuperscript{34,35}. Similarly, another study identified CD45RA\textsuperscript{+}CCR7\textsuperscript{+} HIV-specific CD8\textsuperscript{+} memory T cells functioning as precursors for the other subsets of memory lymphocytes\textsuperscript{36}. Notably, in EBV carriers, 10% of EBV-specific memory CD4\textsuperscript{+} T cells co-express CD45RA and CCR7\textsuperscript{37}. Although formal proof that such cells represented T_{SCM} lymphocytes was missing in these reports, these findings strongly support the hypothesis that T_{SCM} lymphocytes are reproducibly generated after immunization in humans. Differences in the insulting pathogens and tissue environments might impinge on the relative abundance of T_{SCM} generated. Here, we show that T_{SCM} cells have the potential to emerge early after T_N priming, without necessarily passing through an effector stage.
The mechanisms of T-cell memory formation have been actively debated and several models proposed to explain the divergent developmental fates of T-cell progenies. By exploiting the TCR harbored by single T cells infused within the leukapheresis product as surrogate clonal markers, we attempted to evaluate the in vivo differentiation landscapes of single T-cell progenies after transplant in three consecutive patients. Obviously, in the absence of access to secondary lymphoid organs, peripheral or mucosal tissues, constraints typical of studies involving human subjects, we relied on PB as barometer of the entire organism, and tissue-resident progenies of infused lymphocytes could not be included in the present model. With this caveat in mind, we found that upon transfer, discrete T-cell subsets behaved preferentially within a progressive framework of differentiation: infused T_N and T_{SCM} were able to reconstitute the entire spectrum of T-cell memory. Interestingly, up to 30% of originally T_{EM} cells could revert to a T_{CM} phenotype, as reported in non-human primates\textsuperscript{38}. A limited fraction of T_{CM} and T_{EM} reconverted to T_{SCM}, indicating that the T-cell memory compartment is endowed with a certain degree of plasticity in vivo. Our experimental system was characterized by the use of PT-Cy, which we showed to selectively kill the majority of memory T cells proliferating early after transplant. As a result, TCRB sequences originally harbored by LP-T_{SCM}, -T_{CM} and -T_{EM} were found decreased in counts 30 days after transplant. Differently, T_{N} cells were capable of robustly proliferating and differentiating into memory cells. Progenies of individual T_{N} were plastic and had the ability to generate heterogeneous effector and memory populations, comprising T_{SCM}. Notably, the experimental evidence that a third of LP-T_{N} clonotypes traced 30 days after HSCT, were found within the T_{SCM} compartment argues against the hypothesis that T_{SCM} may represent a metastable state within a continuum of phenotypes, but rather reinforce the notion that they represent a stable and relevant memory subset. In our study, only a small proportion of T_{N}-
progenies were traceable in multiple subsets at day 30, while the majority of T<sub>N</sub> clones identified in LP were retrieved in a single memory subclass in vivo. Although the relative contribution of single versus multiple cell fates could not be precisely estimated, our data suggest that the rules of memory T-cell differentiation in humans are at least in part based on population averaging of disparate single-cell behaviors, and are reminiscent of the T-cell dynamics recently described in murine models<sup>6,7,10</sup>.

Altogether, these results highlight T<sub>SCM</sub> lineage as a previously not appreciated important player in the diversification of T<sub>N</sub> cells, and will likely inform strategies aimed at preventing and treating infections<sup>39</sup>, autoimmune diseases<sup>40</sup> and cancer<sup>41</sup>.

**Acknowledgments**

We thank S. R. Riddell (FHCRC, Seattle) for inspiring discussion and continuous advice as second supervisor of N.C. during her PhD, and A. Ambrosi (San Raffaele University) for statistical support. This work was supported by: Italian Ministry of Health (GR07-5 BO, RO10/07-B-1, RF-09-149), Italian Ministry of Research and University (FIRB-IDEAS), Fondazione Cariplo, Italian Association for Cancer Research (AIRC, AIRC 5x1000), EU-FP7 (SUPERSIST, ATTACK2), Regione Piemonte (Direzione Sanità Settore Promozione della Salute e Interventi di Prevenzione Individuale e Collettiva), FIRB Accordi di Programma 2011 (RBAP11T3WB-RNA e nanotecnologie nel controllo dell’immunosoppressione neoplastica sostenuta dal catabolismo degli aminoacidi), Eranet Transcan Haplo_immune.
**Authorship**

N.C. designed the study, conducted laboratory experiments, analyzed data and wrote the manuscript; G.O. performed and analyzed experiments; B.C., V.V. and M.N. assisted in sample processing; M.F., C.T. and S.B. performed bioinformatics analyses; J.P. and R.G. helped in designing the study, provided clinical data and samples, participated to data analyses and interpretation of results; L.V., A.B. and C.Bordignon participated to data discussion; S.M. and F.L. provided clinical samples; L.B. provided leukapheresis samples; F.C. and C.Bonini supervised the study and wrote the paper. N.C. conducted this study as partial fulfillment of her PhD in Cellular and Molecular Biology, and G.O. as partial fulfillment of his PhD in Molecular Medicine, San Raffaele University. C. Bordignon is employee of MolMed S.p.A.; C. Bonini has a research contract with MolMed S.p.A.; all other authors declare no competing interests.

**References**


22. Doubrovina E, Carpenter T, Pankov D, Selvakumar A, Hasan A, O'Reilly RJ. Mapping of novel peptides of WT-1 and presenting HLA alleles that induce epitope-specific HLA-


Figure legends

Figure 1. Donor-derived T\textsubscript{SCM} lymphocytes are selectively enriched early after haploidentical HSCT. (A) Flow cytometric plots of CD8\textsuperscript{+} T-cell phenotype from one representative patient (UPN#15) within the leukapheresis product (LP) and at different time points after HSCT, as labeled. Top row plots depict the expression of CD45RA and CD62L, which allows the identification of T\textsubscript{CM} (CD45RA\textsuperscript{+}CD62L\textsuperscript{+}), T\textsubscript{EM} (CD45RA\textsuperscript{+}CD62L\textsuperscript{-}) and T\textsubscript{EMRA} (CD45RA\textsuperscript{+}CD62L\textsuperscript{-}). Quadrant frequencies are indicated. To discriminate between T\textsubscript{N} and T\textsubscript{SCM}, which are both CD45RA\textsuperscript{+}CD62L\textsuperscript{+}, double positive cells are gated and CD95 expression is shown in bottom row plots. T\textsubscript{SCM} are identified as CD95\textsuperscript{+} while T\textsubscript{N} as CD95\textsuperscript{-}. The percentages of T\textsubscript{SCM} cells gated on CD45RA\textsuperscript{+}CD62L\textsuperscript{+} CD8\textsuperscript{+} T lymphocytes are shown. (B) Summary of CD8\textsuperscript{+} T-cell subset distributions, expressed as percentages on total CD8\textsuperscript{+} T cells, at the indicated time points post HSCT. (C) Comparison of CD8\textsuperscript{+} T-cell subset frequencies at day 8 post HSCT. Each dot represents a single patient. (D) Comparison of the absolute numbers of circulating CD8\textsuperscript{+} T cell subsets at day 8 post HSCT. (E) Comparison of T\textsubscript{SCM} frequencies among LP and day 8 post HSCT. Data are shown as mean values ± S.E.M. of the 20 patients included in the study. (F) Linear regression analysis between CD8\textsuperscript{+} T\textsubscript{SCM} cell counts at day 8 post-HSCT and the serum level of IL-7 measured at day 1 post-HSCT (n=13). Red line denotes the best-fit line of the linear regression analysis, while black dashed lines indicate the 95% confidence interval. The r\textsuperscript{2} and P value of the slope are reported in the upper-right part of the panel.

Figure 2. PT-Cy administration abrogates T-cell proliferation but does not prevent T\textsubscript{SCM} accumulation. (A) Flow cytometry plots depict Ki-67 expression on CD8\textsuperscript{+} T cell subsets from one representative patient (UPN#20) at different time points after HSCT, as labeled. Yellow
arrow denotes PT-Cy administration. (B) Summary of the percentages of proliferating CD8\(^{+}\) T cell subsets at the indicated time points post HSCT. Gray line designates absolute CD8\(^{+}\) T cell counts. Yellow shade indicates PT-Cy administration. (C) Comparison of the percentage of Ki-67\(^{+}\) cells among the different CD8\(^{+}\) T-cell subsets at day 3 post HSCT (prior to PT-Cy administration and in the absence of any immunosuppressive agent). (D) Comparison of the percentages of proliferating cells before PT-Cy administration (day 3) and at the time point when highest T\(_{SCM}\) frequencies were detected (day 8). (E) Flow cytometric plots of ALDH enzymatic activity from one representative patient (UPN#17) of the four tested for ALDH activity. Top row plots show ALDH\(^{+}\) cells at the labeled time points, while bottom row plots show CD34 and CD3 expression on ALDH\(^{+}\) gated cells. (F) Scatter plot depicts the mean percentage of circulating ALDH\(^{+}\) cells at the indicated time points. (G) Scatter plot shows the belonging lineage of ALDH\(^{+}\) cells at the labeled time points: CD34\(^{+}\) cells are indicated with open circles, CD3\(^{+}\) cells with black triangles (\(n=4\)). (H) Percentages of Annexin V\(^{+}\) early apoptotic cells within CD8\(^{+}\) T cell subsets at day 5, 8 and 15 post HSCT. Dashed black line indicates mean percentage of Annexin V\(^{+}\) cells measured in leukapheresis (\(n=3\)). LP indicates leukapheresis product; n.e.: not evaluable.

**Figure 3. CD4\(^{+}\) T\(_{SCM}\) preferentially retain CD31 expression by day 8 post-HSCT.** (A) Box plots depict the differential expression of CD31 on CD4\(^{+}\) T-cell subsets (T\(_{N}\), T\(_{SCM}\), T\(_{CM}\) and T\(_{EM}\)) from healthy subjects (\(n=25\)). (B) Flow cytometry plots of CD31 expression on CD4\(^{+}\) T\(_{N}\), T\(_{SCM}\), T\(_{CM}\) and T\(_{EM}\) cells from one representative patient (UPN#4) at different time points after HSCT, as labeled. (C) Kinetics of representation of T\(_{N}\) and T\(_{SCM}\), expressed as percentage on the total circulating CD4\(^{+}\) cells at the indicated time points post HSCT. T\(_{CM}\) and T\(_{EM}\) kinetics are also
depicted (grey dashed lines). (D) Box plots show CD31 expression on CD4\(^+\) T\(_N\), T\(_{SCM}\), T\(_{CM}\) and T\(_{EM}\) at the labeled time points. Data are shown as average values from the 20 patients included in the study; n.e.: not evaluable.

**Figure 4. Conversion of CD8\(^+\) T\(_N\) into T\(_{SCM}\) can be demonstrated at the antigen-specific level.** (A) Identification of WT1-specific CD8\(^+\) T cells restricted to the host non-shared HLA-A*0101 by dextramer staining of, from top to bottom, donor LP, day 30 peripheral blood and day 30 bone marrow harvested from UPN#4. For each time point analyzed, left plot depicts the percentage of dextramer positive CD8\(^+\) T cells (WT1-dex). Dextramers carrying the same restriction element but loaded with an irrelevant peptide were used to set the gate and distinguish rare antigen-specific cells from background. Central plot shows CD45RA and CD62L expression on dextramer-positive cells (quadrants are set on the polyclonal CD8\(^+\) population). Right plot depicts CD95 expression on CD45RA\(^+\)CD62L\(^+\) cells, when detected (CD95 gate is set on the polyclonal CD8\(^+\) population). (B) Histograms summarizing the quantification and phenotypic characterization of CD8\(^+\) T cells specific for either WT1 (UPN#4, UPN#8, UPN#11 and UPN#13) or PRAME (UPN#2) in 5 patients affected by acute leukemia and with HLA typing suitable for dextramer analysis. For all cytometric dextramer acquisitions no less than 1 x 10\(^6\) cells were stained and no less than 10\(^5\) events on the CD3\(^+\) gate were recorded; n.a.: sample not available; n.e.: population not evaluable.

**Figure 5. In vivo fate mapping by TCR clonotyping uncovers the differentiation landscapes of T\(_N\) cells after transplant.** (A) Graphical representation of the distribution within day 30 memory T-cell subsets of the TCRB sequences harbored by a given T-cell subset from
leukapheresis (LP). Numbers in donut graphs indicate the average percentages, from the three patients analyzed, of TCRB sequences retrieved in the indicated memory subsets 30 days after HSCT. (B) Quantification of the expansion ability of single-cell progenies of T lymphocytes with a given phenotype within LP, independently of the memory T-cell subset in which they were retrieved at day 30 post-HSCT. Expansion is measured as fold change of TCRB sequence count at day 30 relative to LP from one representative patient (UPN#12). (C) Circos plot graphically summarizes the distribution of TCRB sequences originally harbored by LP-T\textsubscript{N} in one representative patient (UPN#12) of the three analyzed. LP-T\textsubscript{N} sequences are depicted in the bottom part of the circle, identified by the green band. From each LP-T\textsubscript{N}-derived TCRB sequence departs a ribbon connecting it to one or more memory subsets at day 30 post-HSCT. Each day 30 memory subset is denoted by a colored band (red for T\textsubscript{SCM}, orange for T\textsubscript{CM} and blue for T\textsubscript{EM/EFF}). Ribbon thickness is proportional to the TCRB sequence count of the given clonotype at day 30 post-HSCT. (D) Left pie chart shows the distribution at day 30 after transplant of the TCRB sequences originally harbored by LP-T\textsubscript{N}. Numbers in the pie indicate the average percentages from the three patients analyzed; right chart zooms on the fraction of clonotypes that were found shared among two or more subsets at day 30 post-HSCT and depicts the mean distribution of all combinations retrieved. (E) Spearman correlation heatmap representing the contribution of individual LP-T\textsubscript{N}-derived clonotypes (row) to the repertoire of T\textsubscript{SCM}, T\textsubscript{CM} and T\textsubscript{EM/EFF} subsets at day 30 post-HSCT (columns). The analysis is performed to LP-T\textsubscript{N} TCRB sequences that are shared by at least two memory T-cell subsets at day 30 post-HSCT. (F) Box plots quantify the expansion, measured as count fold change relative to LP, of LP-T\textsubscript{N} TCRB sequences that were retrieved in memory T-cell subsets at day 30 post-HSCT. (G) Model
scheme of a proposed branched diversification pathway from $T_N$ to $T_{SCM}$, $T_{CM}$ and $T_{EM/EFF}$ subsets upon in vivo transfer following HSCT.
Generation of human memory stem T cells upon haploidentical T-replete hematopoietic stem cell transplantation

Nicoletta Cieri, Giacomo Oliveira, Raffaella Greco, Mattia Forcato, Cristian Taccioli, Beatrice Cianciotti, Veronica Valtolina, Maddalena Noviello, Luca Vago, Attilio Bondanza, Francesca Lunghi, Sarah Marktel, Laura Bellio, Claudio Bordignon, Silvio Bicciato, Jacopo Peccatori, Fabio Ciceri and Chiara Bonini