

ORIGINAL ARTICLE

Quantitative characterization of T-cell repertoire in allogeneic hematopoietic stem cell transplant recipients

PY Yew^{1,5}, H Alachkar^{1,5}, R Yamaguchi², K Kiyotani¹, H Fang¹, KL Yap¹, HT Liu¹, A Wickrema¹, A Artz¹, K van Besien^{1,6}, S Imoto³, S Miyano^{2,3}, MR Bishop¹, W Stock¹ and Y Nakamura^{1,4}

Allogeneic hematopoietic stem cell transplantation (HSCT) is one of curative treatment options for patients with hematologic malignancies. Although GVHD mediated by the donor's T lymphocytes remains the most challenging toxicity of allo-HSCT, graft-versus-leukemia (GVL) effect targeting leukemic cells, has an important role in affecting the overall outcome of patients with AML. Here we comprehensively characterized the TCR repertoire in patients who underwent matched donor or haplo-cord HSCT using next-generation sequencing approach. Our study defines the functional kinetics of each TCRA and TCRB clone, and changes in T-cell diversity (with identification of CDR3 sequences) and the extent of clonal expansion of certain T-cells. Using this approach, our study demonstrates that higher percentage of cord-blood cells at 30 days after transplant was correlated with higher diversity of TCR repertoire, implicating the role of cord-chimerism in enhancing immune recovery. Importantly, we found that GVHD and relapse, exclusive of each other, were correlated with lower TCR repertoire diversity and expansion of certain T-cell clones. Our results highlight novel insights into the balance between GVHD and GVL effect, suggesting that higher diversity early after transplant possibly implies lower risks of both GVHD and relapse following the HSCT transplantation.

Bone Marrow Transplantation (2015) 50, 1227–1234; doi:10.1038/bmt.2015.133; published online 8 June 2015

INTRODUCTION

Allogeneic hematopoietic stem cell transplantation (allo-HSCT) remains one of the most effective therapies for patients with high-risk hematologic malignancies such as AML and myelodysplastic syndromes (MDS). Allo-HSCT using matched related donor (MRD) or HLA-matched unrelated donor (MUD) is the treatment of choice for patients with available matched donors (MD). Umbilical cord blood (UCB) has emerged as an alternative source of hematopoietic stem cells for patients lacking HLA-matched donor, and has lower risk of disease relapse in high-risk leukemia patients.¹ The curative effect of allo-HSCT is largely mediated by the graft-versus-leukemia (GVL) effects, but partially offset by transplant-related mortality and loss of quality of life caused by the GVHD.² Previous studies indicated that patients with GVHD, especially chronic GVHD, are likely to have lower risk of relapse, but increased non-relapse mortality and relatively poor overall survival.^{2,3}

Both GVL and GVHD are mediated by the engrafted T lymphocytes targeting leukemic cells and host tissues, respectively.^{2–4} Before transplantation, HSCT recipients undergo conditioning regimens to eradicate their malignant cells as well as host immune cells that may cause graft rejection. However, since these conditioning regimens include cytotoxic agents and radiation, they cause tissue injury which enhances the induction and secretion of proinflammatory cytokines, and subsequently trigger activation of donor-derived T-cells. These activated donor-derived T-cells that recognize Ag(s) presented by HLA molecules

appear to mediate cytotoxicity against the leukemic cells as well as certain host cells.^{4,5} Several studies established the role of activated donor-derived cytotoxic T-cells in the pathophysiology of acute GVHD.^{6–8} In addition, lower levels of regulatory T-cells have also been suggested to be associated with acute and chronic GVHD.^{9,10}

T-cell receptor (TCR), a heterodimer of TCR alpha and beta chains, or TCR gamma and delta chains, is present on the surface of T-cells. More than 95% of T-cells in our body express heterodimers of TCR alpha and beta chains and the remaining 5% of T-cells express TCR gamma and delta chains.¹¹ The TCR alpha gene is reported to consist of 70 variable (V) exons, 61 joining (J) exons and one constant (C) exon, while the TCR beta gene possesses 60 V exons, 2 diversity (D) exons, 15 J exons and two C exons. In addition, to greatly increase the number of Ags that T-cells can recognize, TCR genes undergo somatic V(D)J recombination that increases the diversity of T-cell repertoire extensively. During the recombination, nucleotides are deleted from V, (D), J exons and/or inserted between a V–J junction for alpha, and V–D and D–J junctions for beta.^{11–13} Hence, the complementarity-determining region 3 (CDR3) of TCRs increases the diversity through the somatic recombination and prepares the immune system to recognize the huge variety of Ags we may be exposed to.¹¹

Although several studies have examined the TCR repertoire in the patients undergoing allo-HSCT, the majority of these studies were limited by their approaches.^{14,15} The use of DNA as a starting materials rather than RNA in some studies limited their ability to

¹Section of Hematology/Oncology, Department of Medicine, The University of Chicago, Chicago, IL, USA; ²Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; ³Health Intelligence Center, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan and ⁴Department of Surgery, The University of Chicago, Chicago, IL, USA. Correspondence: Professor Y Nakamura, Section of Hematology/Oncology, Department of Medicine, The University of Chicago, 900 E. 57th St., Chicago, IL 60637, USA.

E-mail: ynakamura@bsd.uchicago.edu

⁵Both authors contributed equally to this work.

⁶Current address: Hematology/Oncology, Weill Cornell Medical College, New York, NY, USA.

Received 31 December 2014; revised 26 April 2015; accepted 28 April 2015; published online 8 June 2015

address the functional assessment of the T-cell population. Furthermore, TCR-specific multiplex PCR-based assays utilized by these previously published reports restricted these analyses to the known TCRs variable exons and did not allow for a comprehensive unbiased assessment of the TCR repertoire or to identify novel exons that have not been deposited in the public genomic sequence database.^{14–16} Therefore, in order to perform a comprehensive analysis of the TCR repertoire following allo-HSCT, and to examine the kinetic of the TCR profiles and how they relate to the development of GVL or GVHD, we have applied a mRNA-based, 5'RACE (rapid amplification of cDNA ends) PCR method to amplify all the possible combination of TCRA and TCRB using a single set of primers for each of TCRA and TCRB. This was followed by next-generation sequencing (NGS) technology approach which has enabled us to obtain millions of TCR sequences in one experiment and unravel the complexity of T-cell diversity.^{15,16} Here, we describe the detailed TCR repertoire recovery in allo-HSCT patients receiving MD or haploidentical cord (haplo-cord) transplant. We also examined how TCR repertoire and diversity vary by the different source of donor stem cells and how they correlate with clinical variables (relapse or GVHD). TCR repertoire profiling of HSCT patients may lead to a better understanding of the immunological response of patients after HSCT, in order to modulate the GVL-GVHD balance to improve patient's outcome.

MATERIALS AND METHODS

Patients

A total of 21 transplant patients was examined in this study. Twenty patients had AML and one had MDS. The median age of the patients was 56-years old (range: 26–73). The characteristics of patients are summarized in Table 1. Twelve patients received T-cell-depleted MD transplant, while nine patients received haplo-cord transplant. The pre-transplant conditioning regimen consisted of either fludarabine/melphalan/campath or clofarabine/ melphalan/campath in patients who received MD transplant, and fludarabine/melphalan/anti-thymocyte globulin in patients who received haplo-cord transplant. Haplo-cord transplant patients were high-risk AML enrolled on a clinical trial of at the University of Chicago. Three patients had active disease at the time of transplant. The median UCB dose was 1.6×10^7 total nucleated cells (TNC)/kg with HLA cord matching of 4/6 in two patients and 5/6 or 6/6 in seven patients. Among the transplant patients, 11 of them were diagnosed as acute GVHD patients.¹⁷ Peripheral blood mononuclear cells (PBMCs) were obtained from the patients before transplant and at different time points after transplantation. The study was approved by the institutional review board and all patients gave written informed consent on University of Chicago

Table 1. Characteristic of patients

	MD	Haplo-cord
Donor	12	9
HLA identical	5	—
relative		
HLA identical	7	—
unrelated		
Mismatched	—	2
relative + 4/6 UCB		
Mismatched	—	6
relative + 5/6 UCB		
Mismatched	—	1
relative + 6/6 UCB		
Age	42–73	26–67
Conditioning	Flu/Mel/campath or	Flu/Mel/anti-
regimen	Clo/Mel/campath	thymocyte globulin
Acute GVHD	9	2
Relapsed	5	3

Abbreviations: Flu = fludarabine; MD = Matched donor; Mel = melphalan; UCB = Umbilical cord blood.

Institutional Review Board approved treatment trials that included sequential correlative samples (IRB #14736, #14341B and #11300A).

RNA isolation and PCR amplification

Total RNAs were isolated from PBMCs ($\sim 5 \times 10^6$ cells) using RNeasy mini kit (Qiagen, Valencia, CA, USA) and treated with DNase to remove genomic DNA contamination. cDNA was then synthesized using the SMART cDNA library construction kit (Clontech Laboratories, Mountain View, CA, USA), according to the manufacturer's instructions. A common adaptor (SMART IV oligonucleotide) was ligated to the 5'-end of cDNA. Two PCRs were then performed to amplify all the possible combination of TCR alpha and beta from cDNA, using one common forward primer which is designed based on the sequence of SMART IV adaptor and included Ion truncated P1 adaptor sequence (5'-CCTCTCTATGGGCGAGTCGGTATTATCAACGCAGAGTGCCAT-3') and two reverse primers, each specific to the constant region of TCR alpha (5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGCAGGGTCAGGGTTCTGGATA-3') or beta (5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTGATGGCTCAAACACAGC-3') and included the Ion A1 adaptor sequence.¹⁸ The PCR conditions were 94°C for 3 min, followed by 40 cycles of 94°C for 30 s, 65°C for 30 s and 68°C for 1 min.

Template preparation and sequencing

The libraries were further amplified onto the proprietary Ion Sphere particles (ISP) by emulsion PCR using Ion OneTouch2 system and the Ion PGM Template OT2 400 kit (Life Technologies, Carlsbad, CA, USA), according to the manufacturer's instructions. ISPs with templates were then enriched using the Ion OneTouch ES instrument (Life Technologies). Finally, the fully prepared products were subjected to sequencing using the Ion PGM Sequencer (Life Technologies), a 400 bp reading kit and Ion 318 Chip v2.

Sequence analysis

Sequencing reads in FASTQ files were mapped to the reference sequences derived from IMGT/GENE-DB (<http://www.imgt.org>)^{19,20} using Bowtie2 aligner (Version 2.1.0)²¹ The *V*, *D*, *J* genes were designated according to the nomenclature provided by the international ImMunoGeneTics information system (IMGT). A CDR3 region was defined by identifying the second conserved cysteine encoded in the 3' portion of the V segment and the conserved phenylalanine encoded in the 5' portion of the J segment that form the boundaries of the CDR3. The nucleotide sequences between both conserved TCR V cysteine and TCR J phenylalanine were extracted to determine the amino acid sequence of CDR3 region.

Statistical analysis

The inverse Simpson's index was calculated based on the following equation:

$$D_s = \left[\frac{\sum_{i=1}^K n_i(n_i - 1)}{N(N - 1)} \right]^{-1} \quad (1)$$

where *K* is the total number of clonotypes, *n_i* is the number of sequences belonging to the *i*-th type and *N* is the total number of sequences for which clonotypes are determined.

Unpaired Student's *t*-test (two-tailed) was performed for comparison of total proportion of the most abundant ten CDR3 sequences or the diversity index between groups, using GraphPad Prism version 6.0 (La Jolla, CA, USA). A *P*-value of < 0.05 was considered statistically significant.

RESULTS

V(D)J combination and CDR3 sequence analysis

We sequenced the TCR alpha (TCRA) and beta (TCRB) cDNAs synthesized from RNAs that were isolated from PBMCs obtained at several time points from 21 patients who received HSCT. Using a 5' RACE PCR approach, we obtained an average sequence reads of 2 663 492 for TCRA and 1 509 588 for TCRB. The observed sequence reads allowed us to identify 45 TCRAV and 59 TCRBV exons, which represented the majority of the 47 functional V exons in TCRA and 60 functional V exons in TCRB (Supplementary Figures 1a and b), indicating a good coverage of TCR gene by our

cDNA sequencing approach. In addition, interestingly, recombination between TRDV1 and various J exons in TCRA was observed in our samples, indicating that this exon, although defined as a V exon of the TCR delta gene, may also function as a V exon of TCRA (Supplementary Figure 2). We validated it by performing Sanger sequencing and the aligned sequences are shown in Supplementary Figure 3. After defining VJ combinations for TCRA and V(D)J combinations for TCRB, we further defined individual CDR3 sequences using our newly developed algorithm. On average, we were able to identify 28 924 unique CDR3 sequences for TCRA and 26 714 unique CDR3 sequences for TCRB which have open reading frames.

Cord-chimerism correlates with higher TCR repertoire diversity in hapo-cord transplant

We assessed the TCR diversity by calculating the inverse Simpson's diversity index ($1/D_s$) (Supplementary Table S1). As expected, the diversity for TCRA and TCRB was significantly lower in samples obtained after transplant in comparison with pre-transplant samples ($P=0.0054$ and 0.015 for TCRA and TCRB, respectively) (Figures 1a and b). In order to investigate a correlation between the source of donor cells in the haplo-cord transplanted patients and TCR repertoire reconstruction after transplantation, we correlated the diversity of TCRs with the proportion of cord cells in patient's blood at different time points. Patients were dichotomized into two groups according to the median percentage of cord-derived cells at day 30 ($\leq 6\%$ vs $> 6\%$) after haplo-cord transplantation (Figure 2a), and compared the diversity at day 100 between the two groups. We found that TCRA and TCRB of patients in the $> 6\%$ group at day 30 were significantly more diverse at day 100 than those of $\leq 6\%$ group (TCRA; 116.1 ± 17.25 vs 23.87 ± 9.84 and TCRB; 89.59 ± 15.49 vs 21.84 ± 10.51) ($P=0.0017$ and $P=0.0072$ for TCRA and TCRB, respectively) (Figures 2b and c). The same correlation was observed when we examined the diversity on day 100 and percentage of cord-derived cells on day 100 post transplant (data not shown).

However, when we compared the TCR diversity at day 50–100 between patients who underwent MD and those underwent haplo-cord transplant we found no statistical difference in the diversity of TCRA or TCRB ($P=0.27$ and 0.72 , for TCRA and TCRB, respectively, Figures 2d and e). Therefore, we could combine both groups for further correlative analysis.

Expansion of TCR clones in repertoire of patients with GVHD

To examine whether enrichment of certain T-cells may be involved in the development of GVHD, we analyzed changes over time from baseline and sequentially through the development of acute GVHD. On the basis of their V(D)J combination and defined CDR3 sequences (by insertion and deletion of the nucleotides at the rearranged junction), we sorted independent cDNA sequences according to their number of appearance in the sequence reads from the most to least abundant. We demonstrated the 10 most

abundant CDR3 sequences at the time point closest to the date when each patient was diagnosed with acute GVHD. We observed that certain T-cell clones were expanded in each of the GVHD patients and were very low at any other time points. We then compared the combined proportion of top ten CDR3 sequences at the time point closest to acute GVHD diagnosis and multiple time points of non-GVHD patients after HSCT. We found significantly stronger enrichment of the ten most abundant T-cell clones of TCRB in GVHD patients than those with non-GVHD patients ($P=0.0029$). On the other hand we only observed a very modest tendency in the difference in TCRA ($P=0.18$) (Figures 3a and b).

In order to further investigate whether the observed changes in TCR repertoire and diversity were GVHD-related and GVL-independent, we further compared the TCR repertoire between GVHD and non-GVHD patients based on their relapse status. Interestingly, we found that only in non-relapsed patients, the proportions of the 10 most abundant clones (indicating strong clonal expansion) for both TCRA and TCRB were significantly higher in the GVHD group than in the non-GVHD group ($P=0.0031$ and 0.00013 for TCRA and TCRB, respectively; $16.17 \pm 2.09\%$ vs $37.33 \pm 7.73\%$ for TCRA, $14.58 \pm 2.29\%$ vs $35.67 \pm 3.86\%$ for TCRB) (Figures 3c and d). However in the relapsed patient group, we observed no significant difference in the proportion of the top 10 abundant clones between the GVHD and non-GVHD patients (Figures 3e and f). Similarly, we found that diversity of TCRA and TCRB was significantly lower in GVHD patients compared with non-GVHD patients in the non-relapsed group ($P=0.011$ and 0.0064 , for TCRA and TCRB, respectively), (Figures 3g and h), but not in the relapsed group (Figures 3i and j).

Expansion of similar V–J in the GVHD patients with shared HLA alleles

We analyzed the similarity and differences in the V(D)J combination along with CDR3 sequences in the GVHD patients using principal component analysis (PCA) (data not shown). During the analysis, we found enrichment of the same V(D)J combination in both alpha and beta chains in two GVHD patients (Pt# 1 and Pt# 14) who shared five of the six major HLA alleles (*HLA-A*02*, *HLA-B*07*, *HLA-C*07*, *HLA-DR*04*, *HLA-DQ*03* and *HLA-DQ*06*). In these two patients, the TRAV-J (TRAV38-2/DV8-TRAJ30) combination was found at the frequency of $\sim 4.18\%$ and 9.58% , and TRBV15-1-TRBJ2-1 was observed at the frequency of 3.08 and 5.95% (Supplementary Table S2a). Among all the CDR3s found in the TRAV38-2/DV8-TRAJ30 combination, the most abundant one is CAYRVSRRDDKIIF (3.33% for Pt# 1) or CAWRDDKIIF (8.34% for Pt# 14). In the case of TRBV15-1-TRBJ2-1, the most abundant CDR3s is CATSRAYNEQFF for Pt# 1 (2.28%) and CATSRDLYGKNEQFF for Pt# 14 (3.08%) (Supplementary Table S2b).

TCR repertoire diversity correlates with relapse

In order to assess the TCR repertoire in HSCT patients with higher likelihood to have the GVL effect, we examined the correlation of

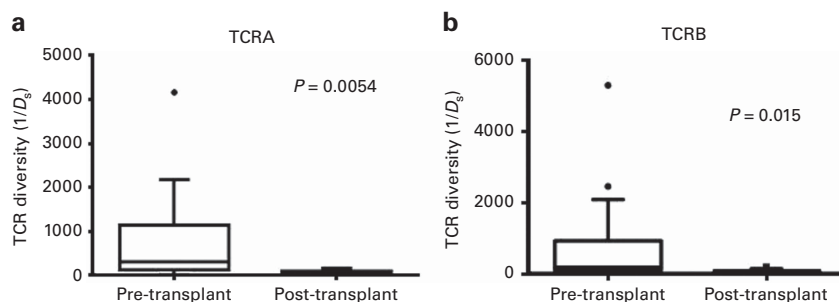


Figure 1. Comparison of diversity index of samples. (a) In all patients, the diversity for TCRA (a) and TCRB (b) before transplantation were compared with that after transplantation.

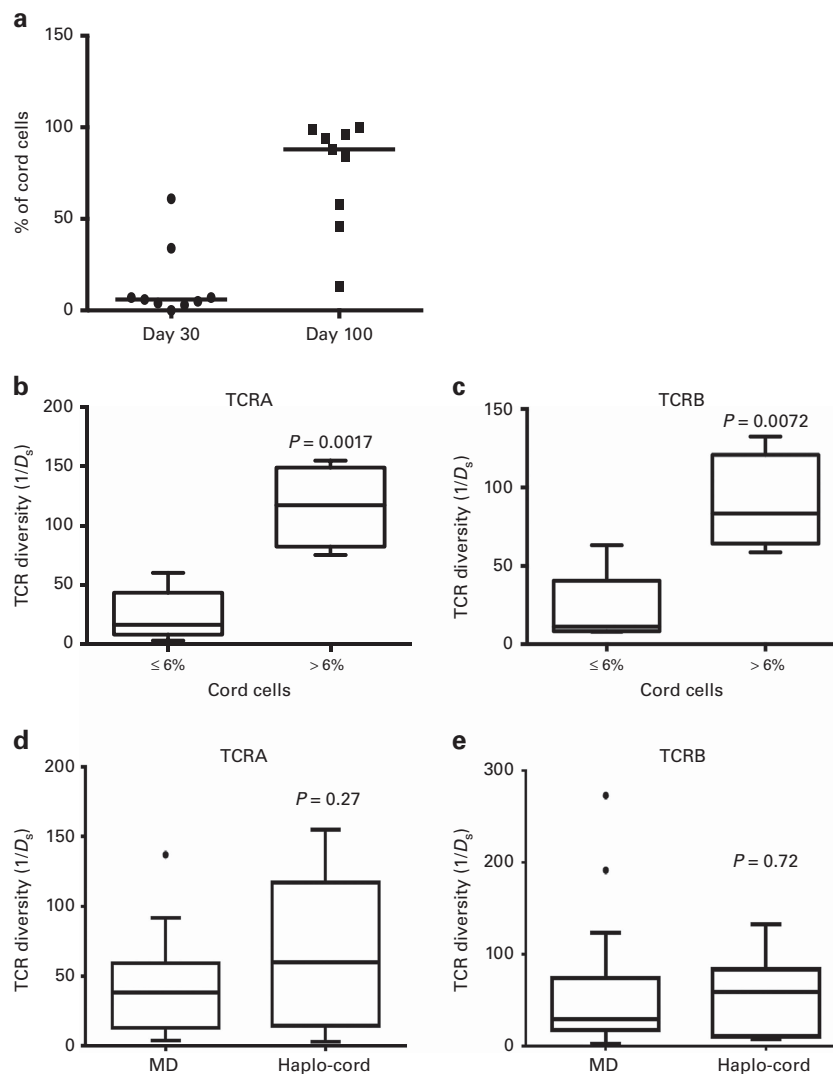


Figure 2. Reconstruction of TCR repertoires observed in the haplo-cord transplant patients at day 100. **(a)** Distribution of patients according to the median of percentage of cord-derived cells at day 30 or day 100 after haplo-cord transplant ($n = 9$). The diversity of TCRA **(b)** and TCRB **(c)** of patients with $> 6\%$ ($n = 4$) cord-derived cells at day 30 were compared with that of patients with $\leq 6\%$ cord-derived cells ($n = 5$). The diversity of TCRA **(d)** and TCRB **(e)** of MD patients were compared with that of haplo-cord patients.

the diversity index of TCR repertoire of samples collected at day 50, 75–100 or 1 year with relapse status. The TCR diversity in relapsed patients was not significantly different from that of non-relapsed patients ($P = 0.82$ and 0.70 , for TCRA and TCRB, respectively, Figures 4a and b). However, in the non-GVHD group, we found that the diversity index was significantly higher in non-relapsed patients compared to that in relapsed patients (85.73 ± 16.59 vs 10.95 ± 4.06 for TCRA; 85.69 ± 14.77 vs 13.25 ± 3.84 for TCRB; $P = 0.042$ for TCRA and 0.029 for TCRB; Figures 4c and d). Consistently, the proportions of the 10 most abundant clones for both TCRA and TCRB were significantly higher in the relapsed patients of the non-GVHD group than in the non-relapsed patients ($P = 0.015$ and 0.017 for TCRA and TCRB, respectively; Supplementary Figures 4a and b).

On the other hand, relapsed patients in the GVHD group showed higher TCRA diversity (82.09 ± 20.46) compared with the non-relapsed patients (29.28 ± 8.10 ; $P = 0.012$), while we only observed a similar trend in the case of TCRB (91.15 ± 38.42 vs 33.77 ± 8.51 for relapsed vs non-relapsed patients; $P = 0.08$, Figures 4e and f). When we compared the proportion of the top ten CDR3 sequences between relapsed and non-relapsed patients in the GVHD group, we observed a modest but not statistically

significant difference in both TCRA and TCRB ($P = 0.11$ and 0.20 for TCRA and TCRB, respectively; Supplementary Figures 4c and d).

DISCUSSION

Experimental models and clinical studies have suggested a link between GVHD and GVL. While GVHD remains the most challenging toxicity of allogeneic transplantation, efforts to mitigate the incidence of GVHD may substantially impact the development of GVL following the transplant. Understanding the balance between GVHD and GVL is particularly important in high-risk AML patients, where great efforts were made to modulate GVHD and enhance the GVL effect after the transplantation in order to reduce the relapse rates and improve clinical outcome for this group of patients. These approaches include manipulation of post-transplantation immune suppression, source of stem cells and cytokine use.^{22,23}

T-cell reconstitution has a major role in the recovery after the transplantation, in long-term resistance to infections, in the development of GVHD and also in the GVL effects. Therefore, to obtain better insight into the interplay between GVHD and GVL in transplanted patients, deeper understanding of the T-cell

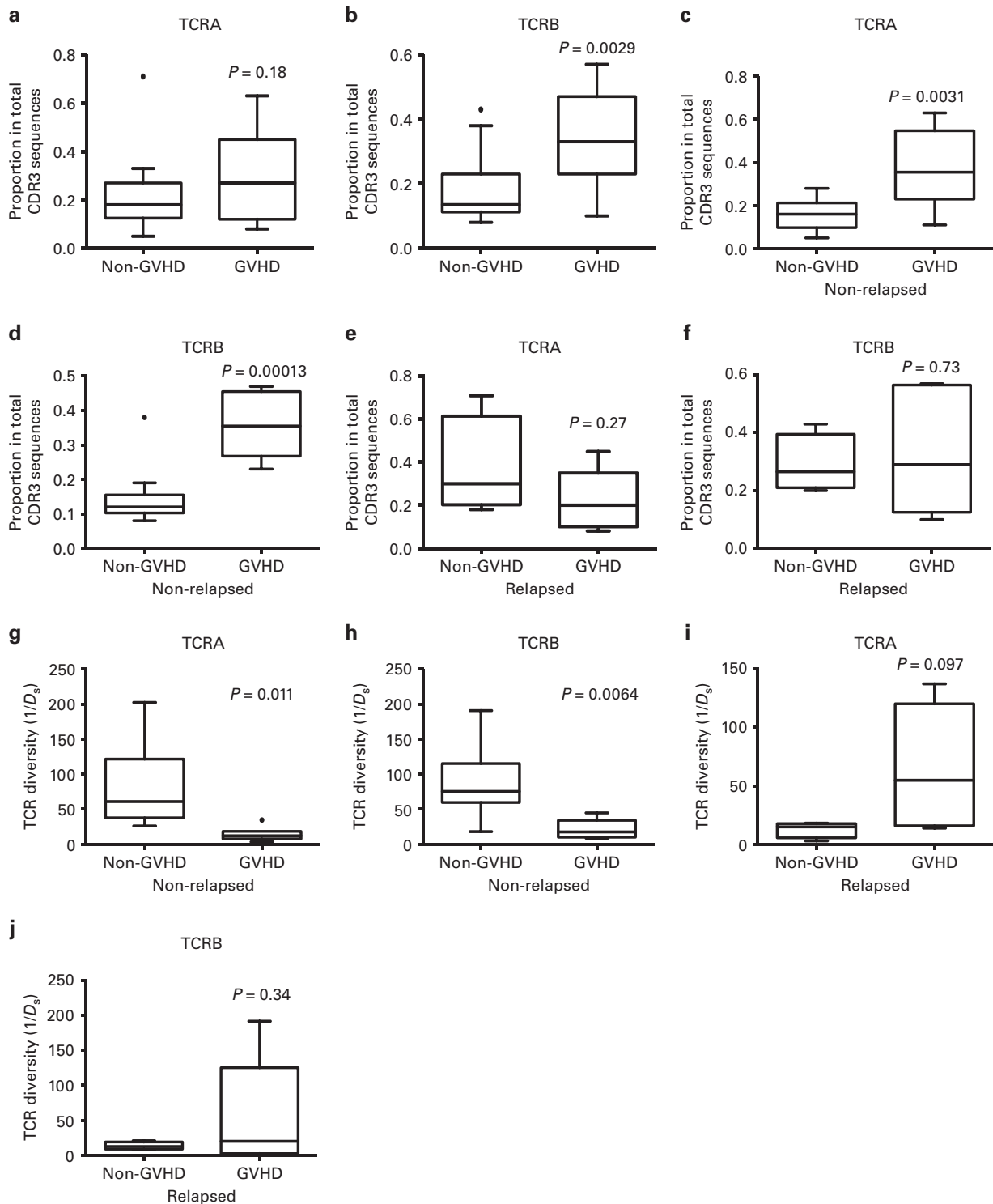


Figure 3. Comparison of the proportion of the top 10 CDR3 sequences and diversity between non-GVHD and GVHD patients. In all patients, the proportion of top 10 CDR3 of TCRA (a) and TCRB (b) of non-GVHD ($n=10$) were compared with that of GVHD patients ($n=11$). In the non-relapsed group, the proportion of top 10 CDR3 of TCRA (c) and TCRB (d) of GVHD patients ($n=6$) were compared with that of non-GVHD patients ($n=7$). In the relapsed group, the proportion of top 10 CDR3 of TCRA (e) and TCRB (f) of GVHD ($n=5$) were compared with that of non-GVHD ($n=3$). In the non-relapsed group, the diversity of TCRA (g) and TCRB (h) of non-GVHD patients ($n=7$) were compared with that of GVHD patients ($n=6$). In the relapsed patients, the diversity of TCRA (i) and TCRB (j) of GVHD ($n=5$) were compared with that of non-GVHD ($n=3$) patients.

reconstitution and their function is essential. GVHD occurs when the donor-derived T-cells recognize some host Ag(s) presented on the HLA molecules at the cell surface as non-self and are activated to effector T-cells which further proliferate and cause damage

to the host cells.^{3,5} T-cells are important in driving the pathophysiology of GVHD, so assessment of T-cell repertoire should help in revealing the molecular mechanism associated with the development of this deleterious disease.

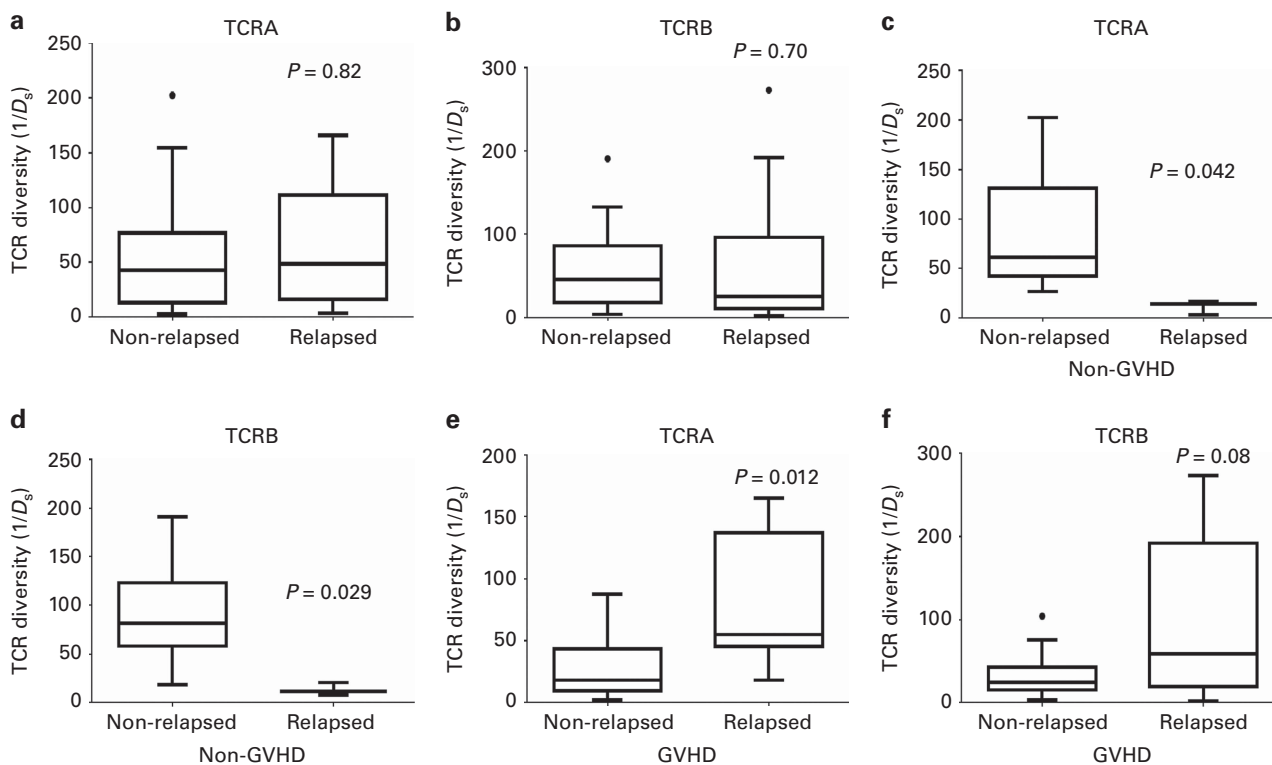


Figure 4. Comparison of the TCR diversity between relapse and non-relapse patients. In all patients, the diversity of TCRA (a) and TCRB (b) of relapsed patients ($n = 7$) were compared with that of non-relapsed patients ($n = 13$). In non-GVHD patients, the diversity of TCRA (c) and TCRB (d) of non-relapsed patients ($n = 7$) were compared with that of relapsed patients ($n = 3$). In GVHD patients, the diversity of TCRA (e) and TCRB (f) of relapsed patients were compared with that of non-relapse patients.

Several reports have examined the TCR repertoire after hematopoietic cell transplantation.^{14,15,24} Most of these studies focused on analyzing the TCRB only; here, we have examined both TCRA and TCRB which are required to form a TCR heterodimer for Ag recognition resulting in higher diversity of the TCR repertoire. Information of both TCRA and TCRB are useful for further functional analysis of TCRs. Additionally, unlike previous attempts that utilize the TCR-specific multiplex PCR to examine TCR repertoire, we applied 5'RACE PCR method to amplify all TCRA and TCRB transcripts using a single set of primers for each of TCRA and TCRB. The use of a single primer set can minimize the amplification bias that occurs by differences in T_m of PCR primers when using exon-specific multiplex primers.¹⁴⁻¹⁶ Furthermore, this approach permits the discovery of potentially novel exons which have not been deposited in the public database. Indeed, we found a transcript generated by recombination between TRDV1 (annotated as a V segment of TCR delta) and various J segments of TCRA in our samples, indicating this TRDV1 may function as one of V segments in the TCR alpha gene. This observation was also recently reported by another study.²⁵

In the present study, we applied the NGS technology to comprehensively analyze TCR sequences at multiple time points in patients who received HSCT, and compared the changes in TCR repertoire in patients who received different sources of stem cells. The repertoire diversity became remarkably low in patients after the transplant. This reduced diversity after HSCT might be related to the conditioning regimen and/or post-transplant immune suppression. On the other hand, higher percentage of cord-blood cells even as early as at 30 days after transplant was correlated with higher TCRA and TCRB diversity thereafter, suggesting that patients with better cord-blood reconstitution might have a better and quick immune reconstitution. The correlation of the TCR diversity with higher percentage of

cord-blood chimerism highlights the critical role of early monitoring of cord-blood cell percentage. This is consistent with previous studies showing that the rapid increase of cord-blood in haplo-cord transplant predicts later myeloid engraftment.²⁶

Our data revealed that certain T-cells were clonally expanded in both GVHD and non-GVHD patients. However, the majority of GVHD patients, particularly in those without relapse, seemed to reveal much higher degree of the expansion of certain T-cells. In most of the GVHD cases, the diversity was significantly lower at the time point that was closest to the day of GVHD diagnosis, compared to that of any time points in the non-GVHD patients. A recent study claimed that higher diversity was observed in the severe GVHD cases, but this might reflect the difference in stem cell sources; the majority of severe GVHD cases in this study received double-unit umbilical cord-blood (DUCB) transplantation.²⁷ Nevertheless, our observation was in accordance with two previous studies in which GVHD revealed lower diversity of T-cell repertoire.^{28,29} It is reasonable to assume that the lower TCR diversity in the GVHD patients was caused by the stronger expansion of certain T-cell populations that might have key roles in the development of GVHD.

On the other hand, interestingly, in the non-GVHD group, non-relapsed patients had a relatively lower extent of enrichment of the 10 most abundant TCR sequences and higher diversity than relapsed patients. Since we observed that the higher diversity was correlated with the lower relapse rates, we speculate that the higher diversity in the TCR sequences may reflect the higher GVL effect and higher possibility for certain anti-leukemia T-cell clones to evolve over time. Therefore, characterization of the TCR repertoire at a later time point after transplantation may allow us to identify T-cell clones associated with the GVL effect. The lower diversity in the relapsed patients is also likely to reflect expansion of certain T-cell populations; as a previous study

suggested that increased levels of regulatory T-cells might suppress the effect of GVL,³⁰ higher levels of regulatory T-cells may confer less GVHD, but could reduce the GVL effect and increase the risk of relapse. On the contrary, we observed the lower diversity of TCR repertoire in the non-relapsed patients compared to relapsed patients in patients who developed GVHD. This may indicate an early evolving of the GVL effect in these patients. This is consistent with previous notions suggesting the lower relapse rate in patients with GVHD.^{2,3}

More interestingly, we observed an expansion of TRAV38-2/DV8-TRAJ30 and TRBV15-1-TRBJ2-1 in two acute GVHD patients who shared all of three major HLA class I alleles (*HLA-A*02*, *HLA-B*07* and *HLA-C*07*) and two of the three major class II loci (*HLA-DR*04*, *HLA-DQ*03* and *HLA-DQ*06*), but the expansion of T-cells having the same V(D)J combination was not observed in the remaining patients. In general, interaction between TCR and a peptide/MHC complex is mediated by the three CDR regions in each of TCR alpha and beta chains.³⁰ CDR1 and CDR2 sequences are considered to be important in the binding to MHC. On the other hand, CDR3 defined as peptide sequences around the V(D)J-rearranged region is considered to be critical to interact with the specific peptide on the MHC molecule. Therefore, Ags recognized by TCRs on T-cells in these two GVHD patients may be different because of different CDR3 sequences. However, several studies suggested that distinct TCRs with different CDR3 (even if the alpha and/or beta chains are different) recognized the same peptide/MHC complex,^{32–34} for example, Davis-Harrison *et al.*³⁴ reported that two different TCRs (different CDR3 sequences in the same TCR beta and different TRAV) could recognize the same Tax/HLA-A2. Hence, it is still possible that although CDR3 sequences are different, T-cells in the two patients might recognize the same Ag on the common MHC molecule.

In summary, our novel quantitative approach using the NGS platform has highlighted that higher cord-chimerism is associated with higher TCR repertoire diversity and that higher diversity early after the transplant possibly implies lower risks of both GVHD and relapse following the HSCT transplantation. Further studies should be performed to examine the role of different T-cell subsets including regulatory T-cells in the pathogenesis of GVHD and also in the relapsed patients. Detailed analysis of the functionally-relevant T-cell repertoire should shed light on better understanding of the significance of the GVL/GVHD interplay in T-cell response/reconstitution which holds essential clinical implications.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

The super-computing resource was provided by Human Genome Center, the Institute of Medical Science, The University of Tokyo (<http://sc.hgc.jp/shirokane.html>).

REFERENCES

- 1 Copelan EA. Hematopoietic stem-cell transplantation. *N Engl J Med* 2006; **354**: 1813–1826.
- 2 van Besien K. Allogeneic transplantation for AML and MDS: GVL versus GVHD and disease recurrence. *Hematology Am Soc Hematol Educ Program* 2013; **2013**: 56–62.
- 3 Gustafsson Jernberg A, Remberger M, Ringdén O, Winiarski J. Graft-versus-leukaemia effect in children: chronic GVHD has a significant impact on relapse and survival. *Bone Marrow Transplant* 2003; **31**: 175–181.
- 4 Ferrara JL, Levine JE, Reddy P, Holler E. Graft-versus-host disease. *Lancet* 2009; **373**: 1550–1561.
- 5 Shlomchik WD. Graft-versus-host-disease. *Nat Rev Immunol* 2007; **7**: 340–352.
- 6 Ichiki Y, Bowlus CL, Shimoda S, Ishibashi H, Vierling JM, Gershwin ME. T cell immunity and graft versus host disease (GVHD). *Autoimmun Rev* 2006; **5**: 1–9.

- 7 Schmaltz C, Alpdogan O, Muriglan SJ, Kappel BJ, Rotolo JA, Ricchetti ET *et al*. Donor T cell-derived TNF is required for graft versus host disease and graft versus tumor activity after bone marrow transplantation. *Blood* 2003; **101**: 2440–2445.
- 8 Kataoka Y, Iwasaki T, Kuroiwa T, Seto Y, Iwata N, Hashimoto N *et al*. The role of donor T cells for target organ injuries in acute and chronic graft versus host disease. *Immunology* 2001; **103**: 310–318.
- 9 Dong S, Maiella S, Xhaard A, Pang Y, Wenyandy L, Larghero J *et al*. Multiparameter single-cell profiling of human CD4+FOXP3+ regulatory T-cell populations in homeostatic conditions and during graft-versus-host disease. *Blood* 2013; **122**: 1802–1812.
- 10 Rezvani K, Mielke S, Ahmadzadeh M, Kilic Y, Savani BN, Zeilich J *et al*. High donor FOXP3-positive regulatory T-cell (Treg) content is associated with a low risk of GVHD following HLA-matched allogeneic SCT. *Blood* 2006; **108**: 1291–1297.
- 11 Janeway CA, Travers P, Walport M, Shlomchik MJ. *Immunobiology: The Immune System in Health and Disease* 5th edition 2005. Glossary, Garland Science, New York.
- 12 Scaviner D, Lefranc M-P. The human T cell receptor alpha variable (TRAV) genes. *Exp Clin Immunogenet* 2000; **17**: 83–96.
- 13 Folch G, Lefranc M-P. The human T cell receptor beta variable (TRBV) genes. *Exp Clin Immunogenet* 2000; **17**: 42–54.
- 14 Meier J, Roberts C, Avent K, Hazlett A, Berrie J, Payne K *et al*. Fractal organization of the human T cell repertoire in health and after stem cell transplantation. *Biol Blood Marrow Transplant* 2013; **19**: 366–377.
- 15 Chapman M, Warren EH, Wu CJ. Applications of next-generation sequencing to blood and marrow transplantation. *Biol Blood Marrow Transplant* 2012; **18**(Suppl 1): S151–S160.
- 16 Woodsworth DJ, Castellarin M, Holt RA. Sequence analysis of T-cell repertoires in health and disease. *Genome Med* 2013; **5**: 98.
- 17 Przepiorka D, Weisdorf D, Martin P, Klingemann HG, Beatty P, Hows J *et al*. 1994 Consensus Conference on Acute GVHD Grading. *Bone Marrow Transplant* 1995; **15**: 825–828.
- 18 Fang H, Yamaguchi R, Liu X, Daigo Y, Yew PY, Tanikawa C *et al*. Quantitative T cell repertoire analysis by deep cDNA sequencing of T cell receptor α and β chains using next-generation sequencing (NGS). *Oncoimmunology* 2015; **3**: e968467.
- 19 Giudicelli V, Denys C, Lefranc M-P. IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Res* 2005; **33**: D256–D261.
- 20 Lefranc MP, Clement O, Kaas Q, Duprat E, Chastellan P, Coelho I *et al*. IMGT, the International ImmunoGeneTics Information System. *Nucleic Acids Res* 2005; **33**: D593–D597.
- 21 Langmead B, Steven LS. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012; **9**: 357–359.
- 22 Gale RP, Horowitz MM, Ash RC, Champlin RE, Goldman JM, Rimm AA *et al*. Identical-twin bone marrow transplants for leukemia. *Ann Intern Med* 1994; **120**: 646–652.
- 23 Hashem H, Lazarus HM. Double umbilical cord blood transplantation: relevance of persistent mixed-unit chimerism. *Biol Blood Marrow Transplant* 2015; **21**: 612–619.
- 24 Kanda J, Chiou LW, Szabolcs P, Sempowski GD, Rizzieri DA, Long GD *et al*. Immune recovery in adult patients after myeloablative dual umbilical cord blood, matched sibling, and matched unrelated donor hematopoietic cell transplantation. *Biol Blood Marrow Transplant* 2012; **18**: 1664–1676.
- 25 Liu P, Liu D, Yang X, Gao J, Chen Y, Xiao X *et al*. Characterization of human α TCR repertoire and discovery of D-D fusion in TCR β chains. *Protein Cell* 2014; **5**: 603–615.
- 26 Kwon M, Balsalobre P, Serrano D, Perez Corral A, Buño I, Anguita J *et al*. Single cord blood combined with HLA-mismatched third party donor cells: comparable results to matched unrelated donor transplantation in high-risk patients with hematologic disorders. *Biol Blood Marrow Transplant* 2013; **19**: 143–149.
- 27 van Heijst JW, Ceberio I, Lipuma LB, Samilo DW, Wasilewski GD, Gonzales AM *et al*. Quantitative assessment of T cell repertoire recovery after hematopoietic stem cell transplantation. *Nat Med* 2013; **19**: 372–377.
- 28 Liu C, He M, Rooney B, Kepler TB, Chao NJ. Longitudinal analysis of T-cell receptor variable beta chain repertoire in patients with acute graft-versus-host disease after allogeneic stem cell transplantation. *Biol Blood Marrow Transplant* 2006; **12**: 335–345.
- 29 Du JW, Gu JY, Liu J, Cen XN, Zhang Y, Ou Y *et al*. TCR spectratyping revealed T lymphocytes associated with graft-versus-host disease after allogeneic hematopoietic stem cell transplantation. *Leuk Lymphoma* 2007; **48**: 1618–1627.
- 30 Barrett AJ, Battiwalla M. Relapse after allogeneic stem cell transplantation. *Expert Rev Hematol* 2010; **3**: 429–441.
- 31 Roomp K, Domingues FS. Predicting interactions between T cell receptors and MHC-peptide complexes. *Mol Immunol* 2011; **48**: 553–562.
- 32 Ding YH, Smith KJ, Garboczi DN, Utz U, Biddison WE, Wiley DC. Two human T cell receptors bind in a similar diagonal mode to the HLA-A2/Tax peptide complex using different TCR amino acids. *Immunity* 1998; **8**: 403–411.

- 33 Cristina C, David NP, Rafick-Pierre S, Francois D. Highly Biased CDR3 Usage in Restricted Sets of β Chain Variable Regions During Viral Superantigen 9 Response. *J Exp Med* 1998; **187**: 253–258.
- 34 Davis-Harrison RL, Armstrong KM, Baker BM. Two different T cell receptors use different thermodynamic strategies to recognize the same peptide/MHC ligand. *J Mol Biol* 2005; **346**: 533–550.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Supplementary Information accompanies this paper on Bone Marrow Transplantation website (<http://www.nature.com/bmt>)