## Supplementary Figures and Tables

### Figure S1

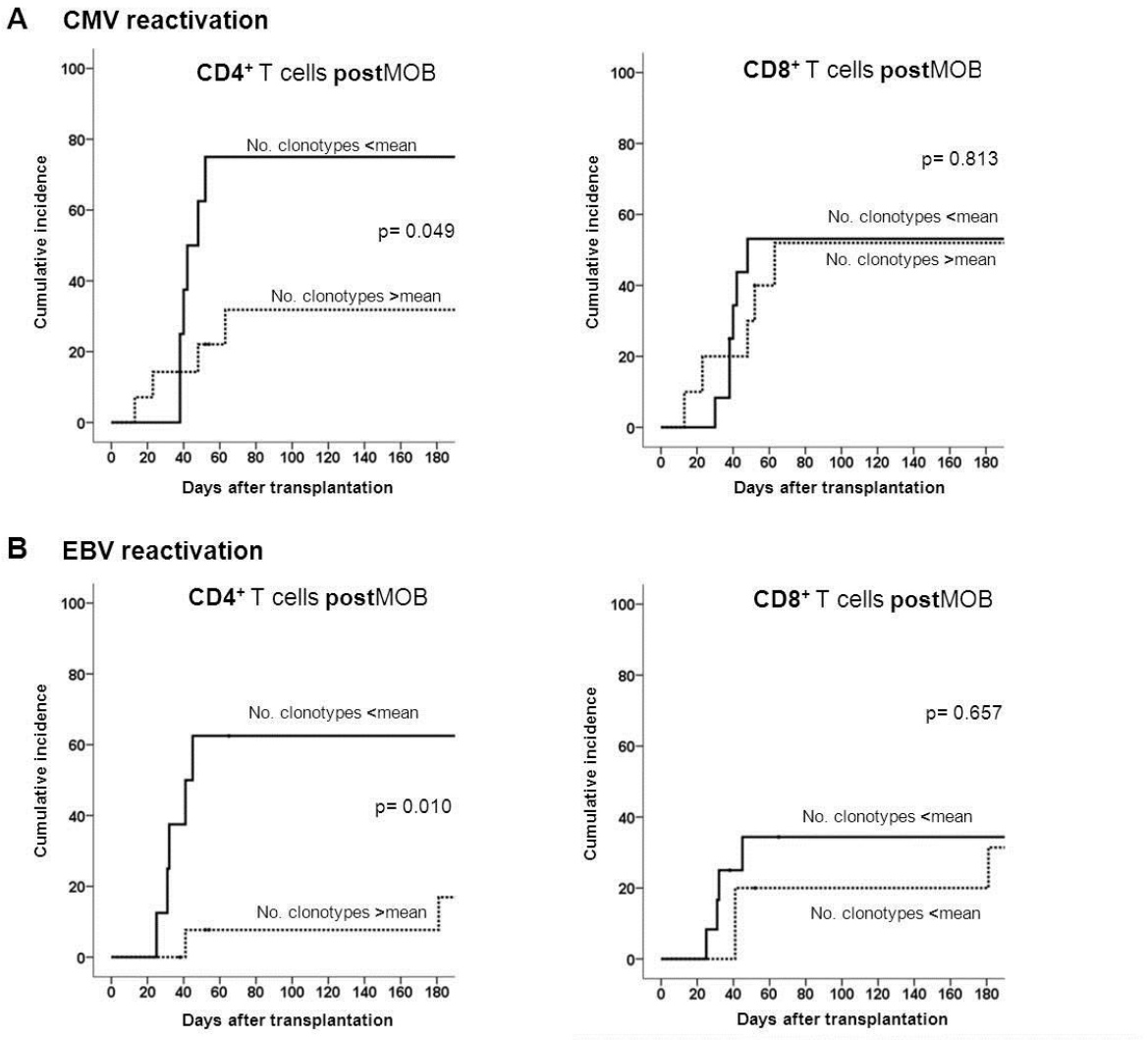



**Fig S1. Cumulative incidence of viral reactivation after allogeneic stem cell transplantation.** Reactivation was monitored weekly by PCR method. Number of clonotypes in donors after G-CSF mobilization were grouped in below and above mean value.

**Figure S2: Control individuals**

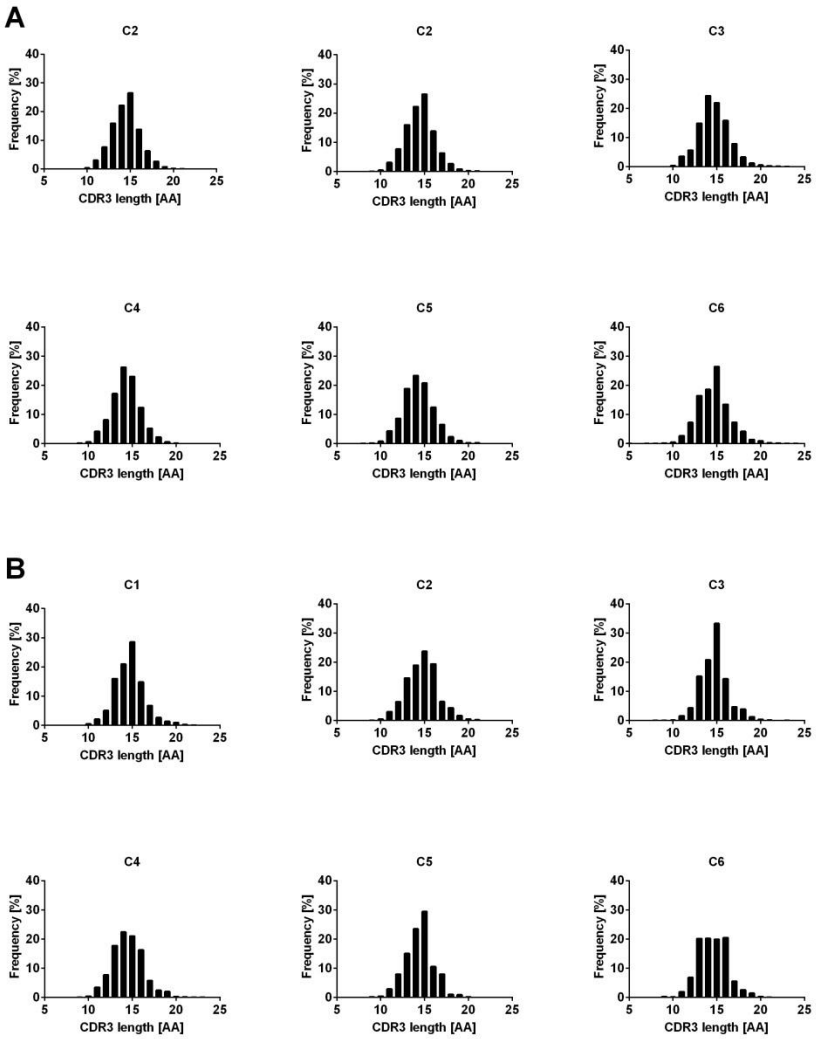
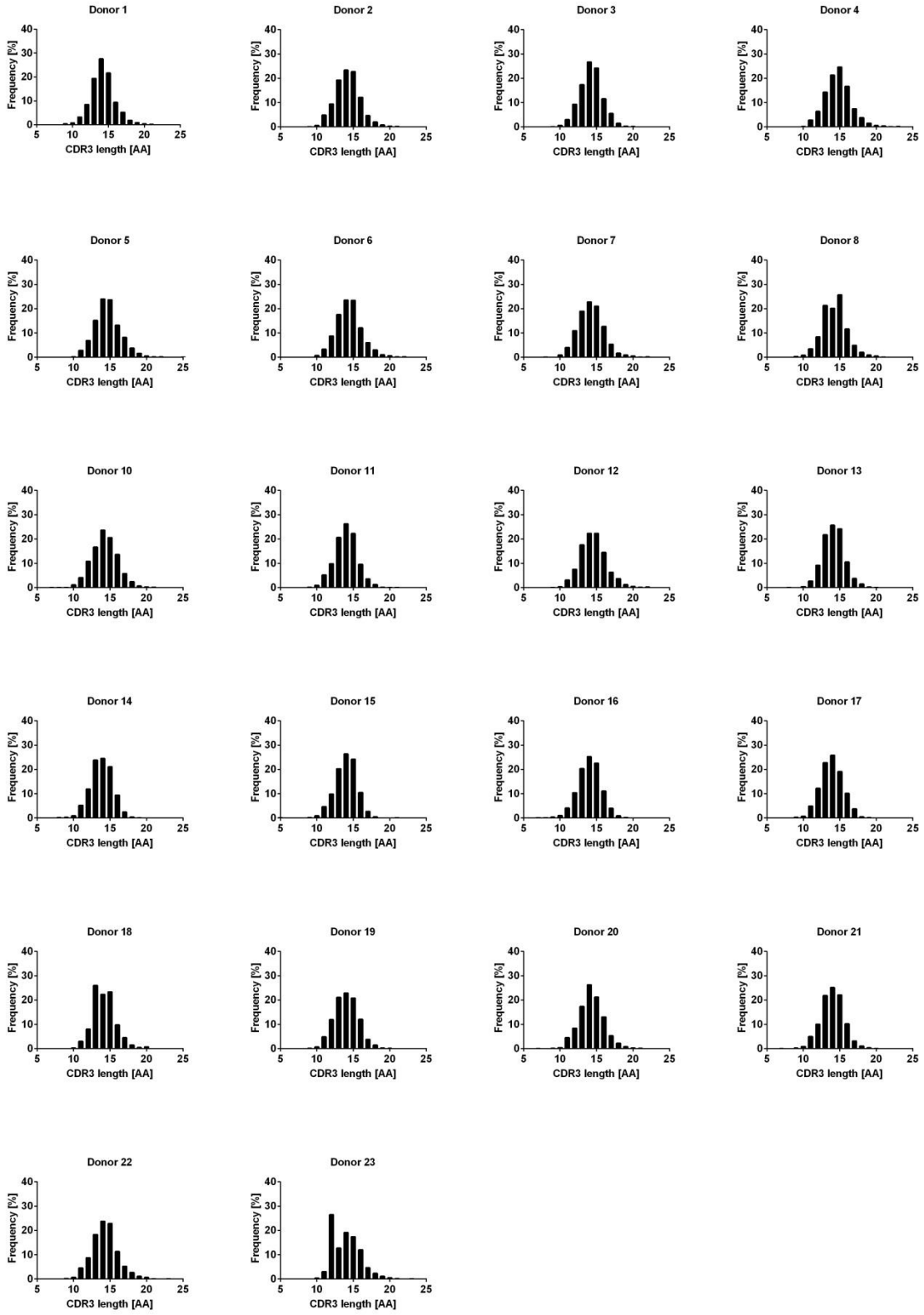
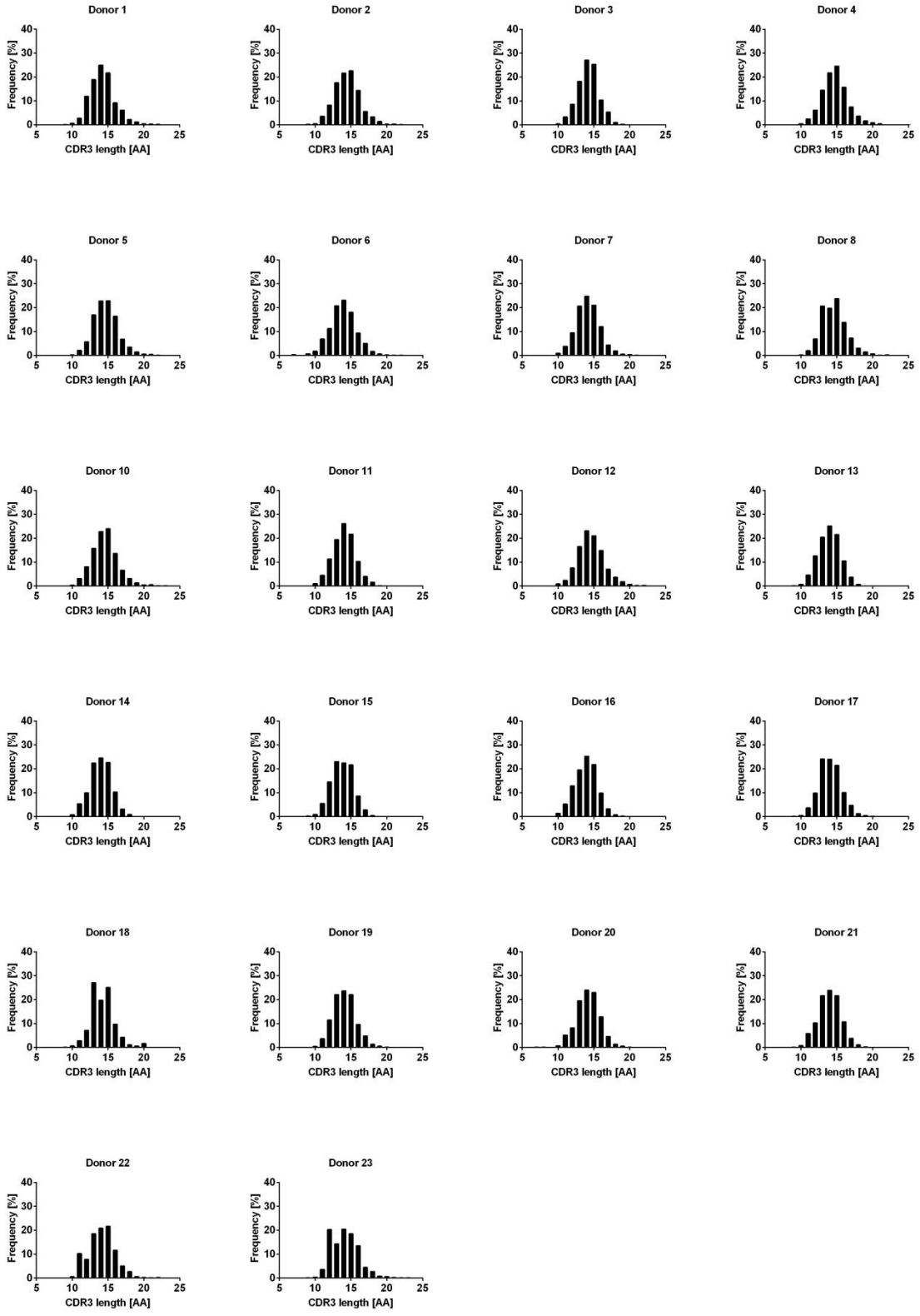



**Fig S2. CDR3 length distribution in healthy control individuals.** HTS derived CDR3 length distribution of (**A**) $CD4^+$ or (**B**) $CD8^+$ TCRβ sequences for every control donor ('C1' – 'C6'. TCRβ CDR3 size is defined as all amino acids (AA) starting from the conserved 5' cysteine in the V segment and ending at the conserved 3' phenylalanine in the J segment.

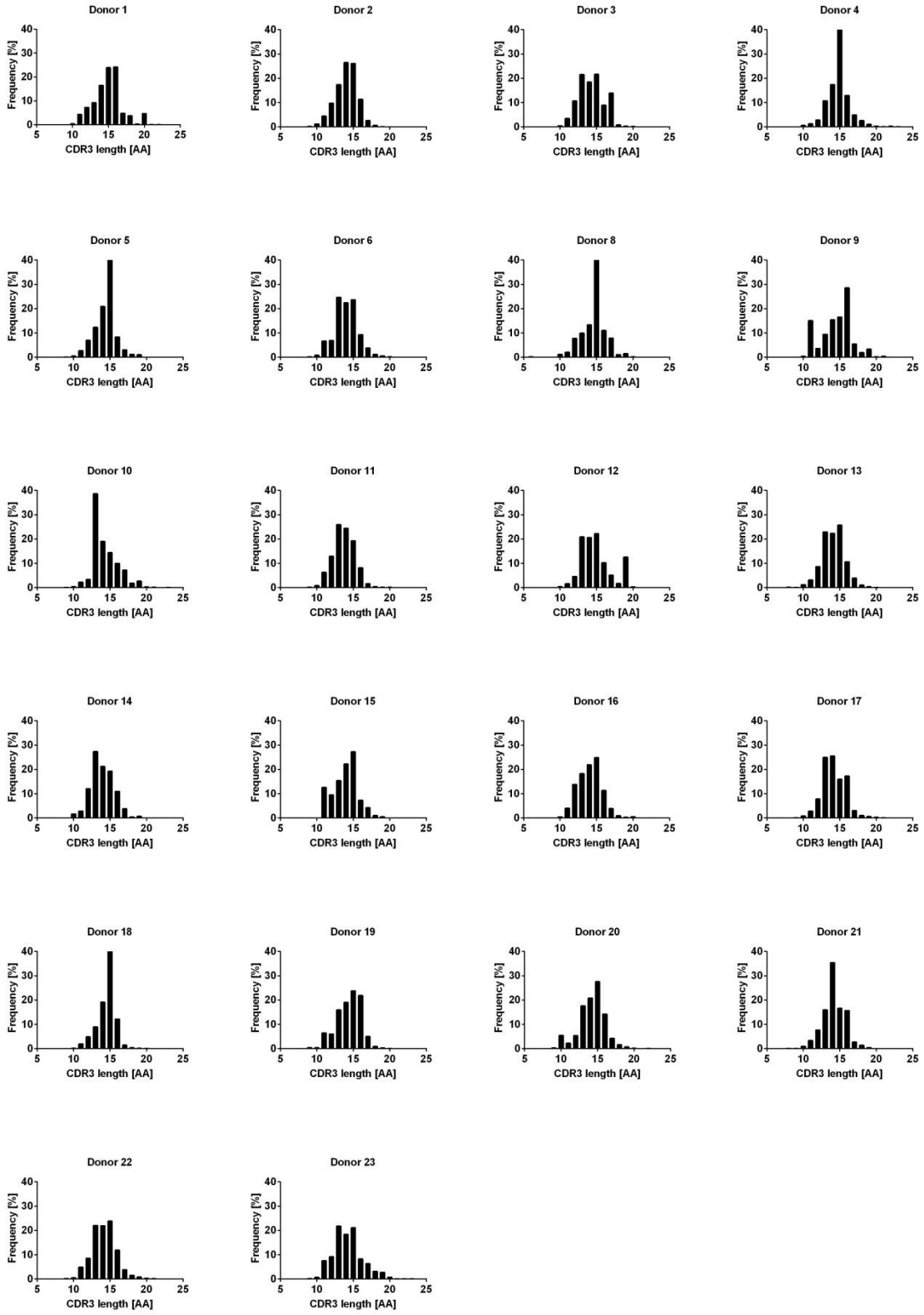
## Figure S3 (A): CD4$^{+}$ T cells preMOB



## Figure S3 (B): CD4$^{+}$ T cells postMOB



## Figure S3 (C): CD8$^{+}$ T cells preMOB



## Figure S3 (D): CD8$^{+}$ T cells postMOB

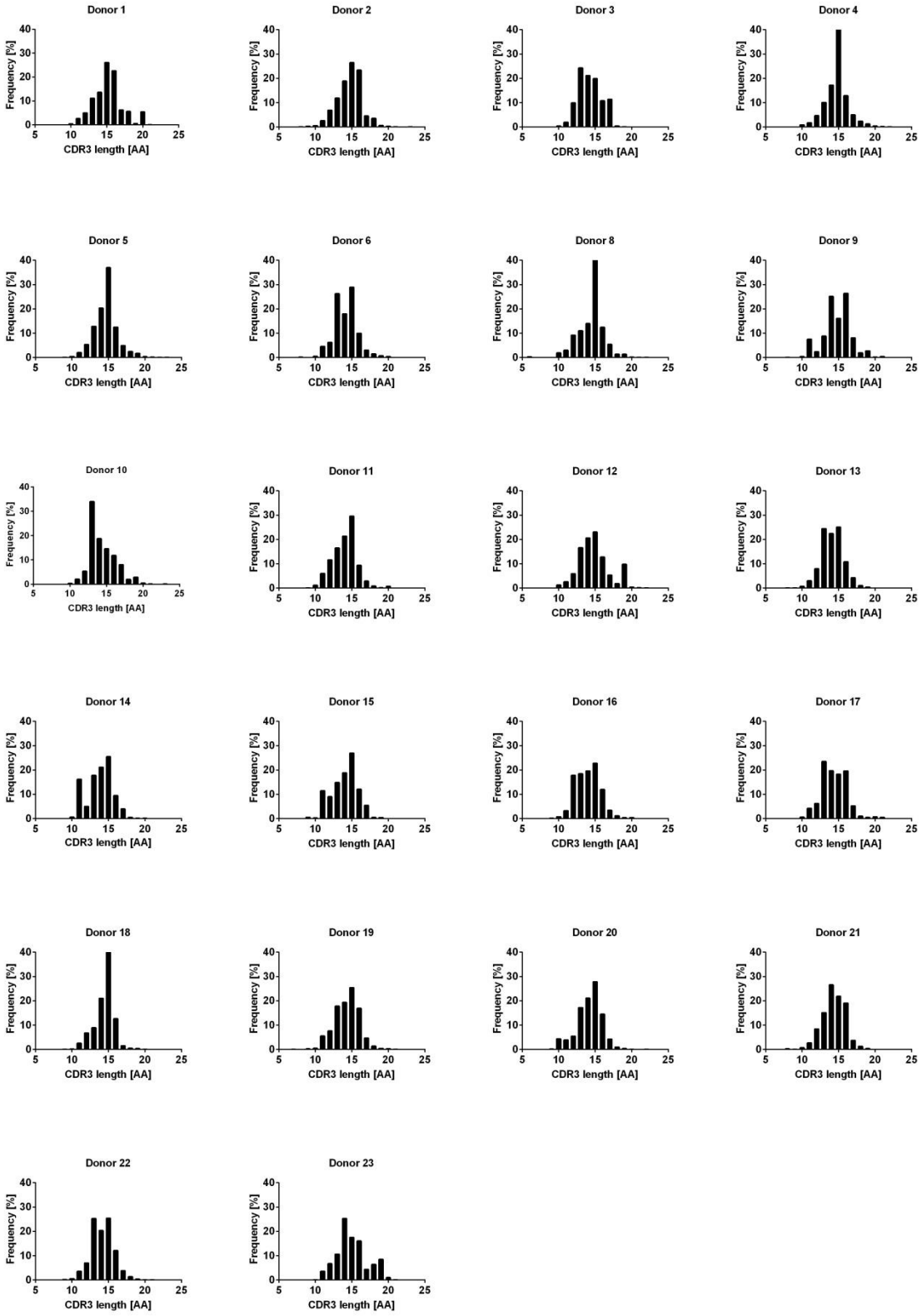


**Fig S3. HTS derived CDR3 length distribution of TCRβ sequences in G-CSF mobilized stem cell donors.** HTS derived CDR3 length distribution of $CD4^+$ TCRβ sequences before (pre, **A**) or after (post, **B**) G-CSF induced stem cell mobilization as well as the HTS derived CDR3 length distribution of $CD8^+$ TCRβ sequences pre (**C**) or post (**D**) G-CSF administration for every stem cell donor. TCRβ CDR3 size is defined as all amino acids (AA) starting from the conserved 5' cysteine in the V segment and ending at the conserved 3' phenylalanine in the J segment.

## Figure S4

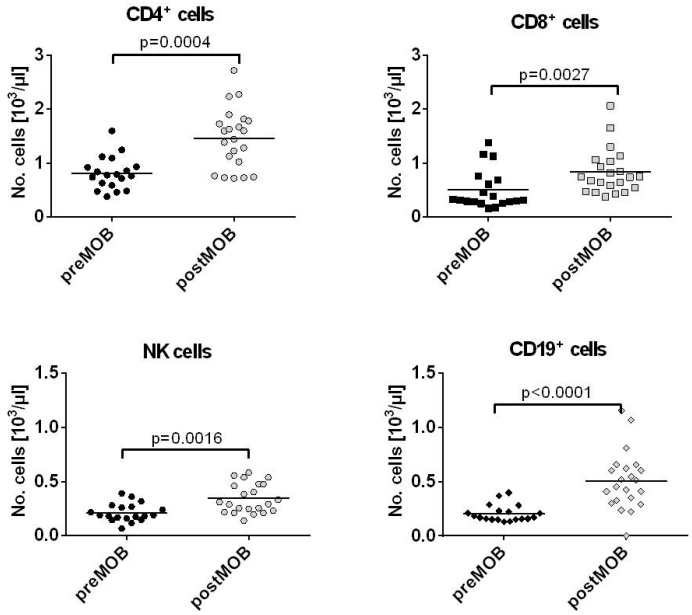



**Fig S4. Cell counts in peripheral blood of donors before and after G-CSF administration.** Diversity dot blots of cell counts of peripheral blood lymphocyte subsets before (preMOB) stem cell mobilization via G-CSF and at the day of apheresis (postMOB). Leukocytes were gated as $CD45^{+}$ and lymphocytes as $CD45^{high}CD14^{-}$ cells. Within the lymphocytic population, T cells were determined as $CD3^{+}$, B cells as $CD19^{+}$, NK cells as $CD56^{+}$ $CD3^{-}$ cell populations. T-cell subpopulations were analyzed upon CD4 and CD8 expression. Cell counts/µl whole blood were calculated based on the number of beads and the sample volume in TruCount tubes (BD Bioscience). The numbers of all lymphocytic subsets increased significantly after G-CSF administration (Mann-Whitney *U* test).

**Figure S5**

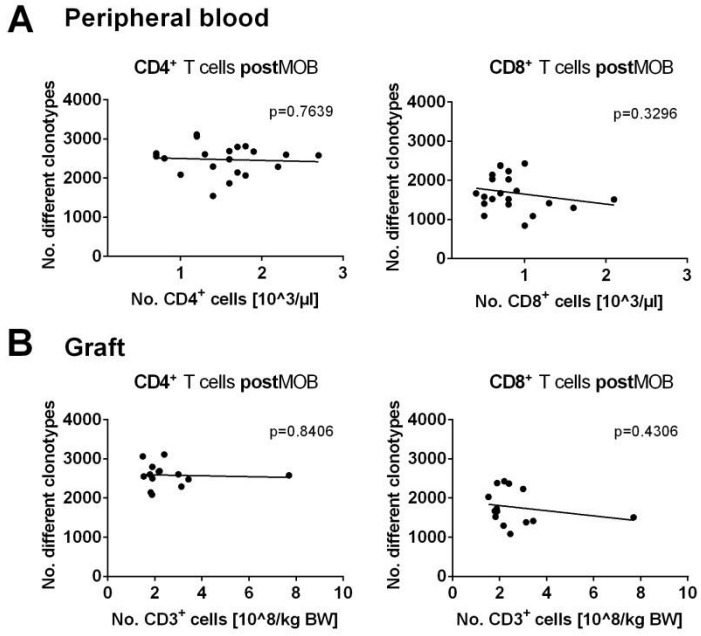



**Fig S5. TCR diversity in correlation with cell numbers of the donor peripheral blood or the stem cell graft.** Scatter plots of $CD4^+$ and $CD8^+$ T-cell diversity and (**A**) absolute T-cell counts ($CD4^+$ and $CD8^+$ T cells) in peripheral blood of stem cell donors post G-CSF mobilization and (**B**) cell counts of $CD3^+$ cells (per kg recipient bodyweight) in the graft. There was no correlation of TCRβ diversity with peripheral blood cell counts.

## Figure S6

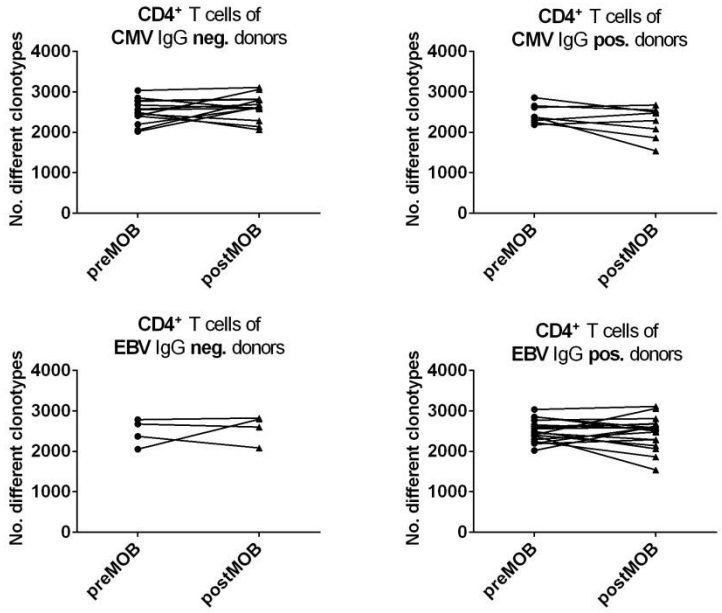



**Fig S6. CD4$^+$ TCR repertoire diversity of CMV/EBV seropositive or seronegative donors before and after G-CSF mobilization.** Diversity dot blots of CMV and/or EBV seronegative (left panel) and seropositive donors (right panel). Changes in CD4$^+$ T-cell diversity after G-CSF treatment are independent of the serostatus of the donor.

**Figure S7**

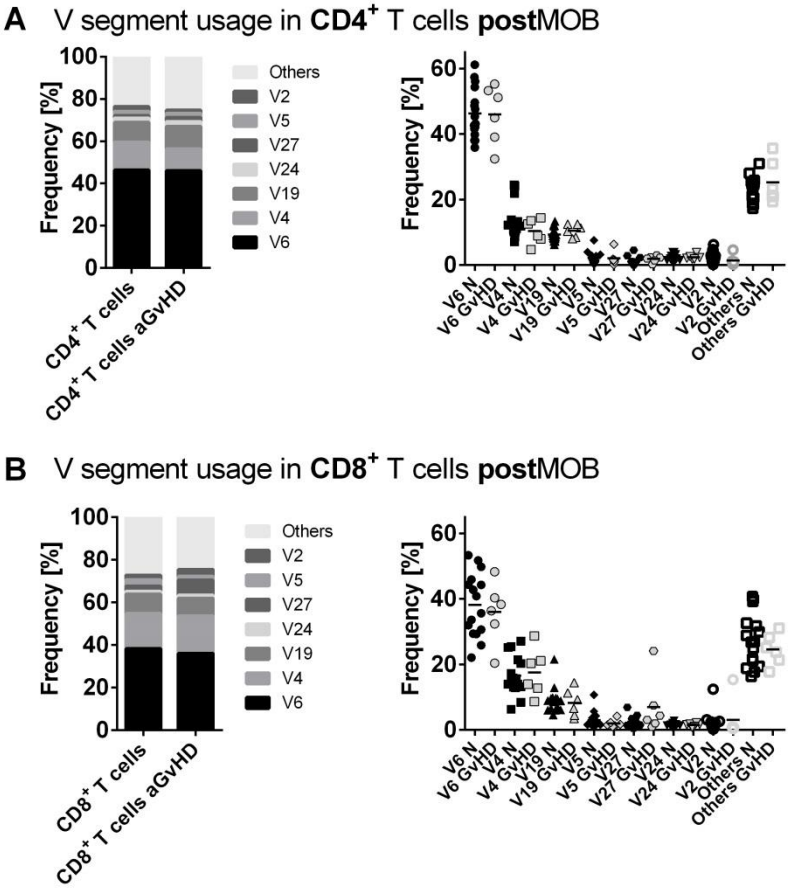



**Fig S7. Vβ subgroup usage in donors whose recipients later suffered from GvHD.** Mean frequencies of Vβ subgroup usage of TCRs from $CD4^+$ (**A**) or $CD8^+$ (**B**) T cells in donors segregated by appearance of aGvHD in recipients. Seven of the most frequent Vβ subgroups were illustrated individually in the bar plot (left panel). The remaining subgroups were compiled. The nomenclature according to the international ImMunoGeneTics information system (IMGT) was used. The right panel shows the individual frequencies of all donors. No significant differences could be shown between these two groups.

**Figure S8**

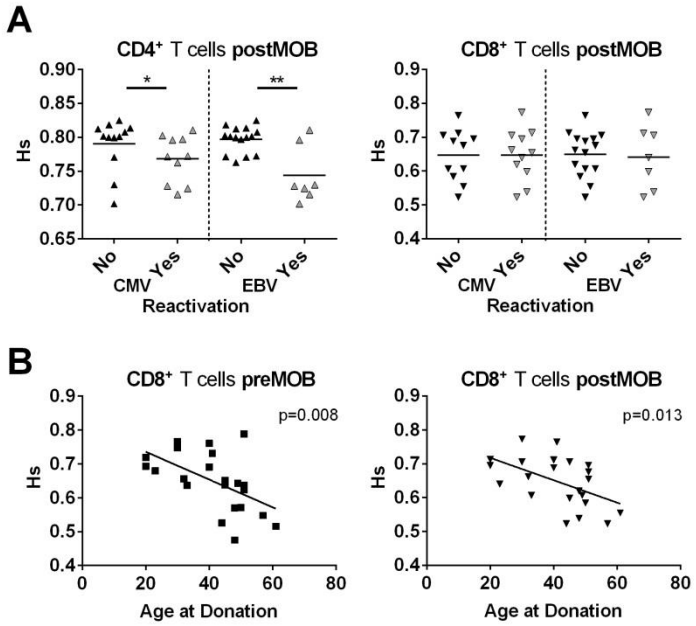



**Fig S8. Donor TCRβ analyses using Shannon entropy ($H_S$).** (**A**) Diversity dot plots of $CD4^+$ and $CD8^+$ T-cell preparation from G-CSF mobilized grouped according to reactivation of CMV or EBV. Donors of patients without CMV and/or EBV reactivation showed significant higher diversity in the $CD4^+$ T-cell compartment post G-CSF mobilization (*p=0.034; **p=0.003). In $CD8^+$ T cells, no significant difference could be detected. (**B**) Scatter plots of $CD8^+$ T-cell diversity and the age for stem cell donors pre and post G-CSF mobilization. The diversity of $CD8^+$ T cells decreases with increasing donor age irrespective of G-CSF mobilization (preMOB: r=-0.55; p=0.008; postMOB: r=-0.52; p=0.013; n=22).

**Table S1.** Characteristics of 23 stem cell recipients.

| Recipient | | | | | | | | | | | | | Graft | Donor | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No. | Sex | Age | Dis. | Conditioning regime | IgG | | acute GvHD | Reactivation | | Relapse | Chimerism | | $CD3^+$ | $CD4^+$ | $CD8^+$ |
| | | | | | CMV | EBV | | CMV (max. copies) | EBV (max. copies) | | Day | Donor % | [10^8/kg BW] | [10^3/µl PB] | |
| 1 | M | 60 | MM | RIC / FBM | pos. | pos. | Yes | No | Yes (70000) | No | 30 | 99.9 | N.A. | 2.2 | 1.0 |
| 2 | M | 62 | MM | RIC / FBM | neg. | pos. | Yes | No | No | No | 60 | 100 | 2.4 | 1.2 | 0.7 |
| 3 | M | 60 | MM | RIC / FBM | pos. | pos. | Yes | Yes (42000) | No | No | 38 | 99.9 | 3.0 | 1.3 | 0.8 |
| 4 | M | 27 | AML | MAB / TreoFlu | neg. | pos. | No | No | No | Yes | 60 | 99.3 | N.A. | N.A. | N.A. |
| 5 | F | 57 | CLL | RIC / TTFluCy | neg. | pos. | No | No | No | Yes | 60 | 99.7 | N.A. | 2.3 | 0.6 |
| 6 | M | 50 | CML | MAB /TreoFlu | pos. | pos. | Yes | Yes (23000) | Yes (430000) | Yes | 30 | 100 | N.A. | 1.4 | 0.5 |
| 7 | M | 57 | AML | RIC / FBM | neg. | pos. | Yes | No | No | No | 30 | 100 | 1.5 | 1.2 | 0.4 |
| 8 | F | 47 | AML | RIC / FBM | pos. | pos. | No | No | No | No | 30 | 100 | 2.2 | 1.9 | 1.6 |
| 9 | M | 67 | MDS | RIC / FBM | pos. | pos. | Yes | Yes (24000) | No | No | 35 | 100 | 2.5 | 0.7 | 1.1 |
| 10 | M | 57 | ALL | RIC / BusFlu | neg. | pos. | Yes | No | No | No | 37 | 100 | N.A. | 0.7 | 0.5 |
| 11 | M | 56 | OMF | RIC / BusFlu | neg. | N.A. | Yes | No | No | No | 30 | 81.6 | N.A. | 1.8 | 0.8 |
| 12 | F | 51 | AML | RIC / FBM | pos. | pos. | Yes | Yes (12000) | Yes (320000) | Yes | 30 | 100 | 1.5 | 0.7 | 0.6 |
| 13 | M | 66 | AML | RIC / FBM | pos. | pos. | No | Yes (3800) | Yes (28000) | Yes | 40 | 100 | 2.2 | 1.6 | 1.0 |
| 14 | M | 60 | MDS | RIC / FBM | pos. | pos. | Yes | Yes (75000) | No | No | 36 | 100 | 1.8 | 1.7 | 0.6 |
| 15 | M | 56 | AML | RIC / FBM | pos. | pos. | No | Yes (26000) | No | No | 39 | 93 | N.A. | 1.8 | 0.5 |
| 16 | M | 59 | MDS | RIC / FBM | pos. | pos. | Yes | Yes (140000) | No | No | 30 | 100 | 1.9 | 1.0 | 0.9 |
| 17 | M | 46 | AML | RIC / FBM | pos. | pos. | Yes | Yes (23000) | No | No | 30 | 100 | 1.9 | 0.8 | 0.4 |
| 18 | M | 52 | MDS | RIC / FBM | pos. | pos. | Yes | Yes (30000) | Yes (42000) | No | 30 | 100 | 3.1 | 1.4 | 0.8 |
| 19 | F | 33 | AML | RIC / FBM | neg. | pos. | Yes | No | No | No | N.A. | N.A. | 1.8 | 0.7 | 0.7 |
| 20 | M | 22 | AML | RIC / FBM | neg. | neg. | Yes | No | No | No | 29 | 100 | 1.9 | 1.7 | 0.7 |
| 21 | M | 69 | AML | RIC / FBM | neg. | pos. | Yes | No | No | No | 29 | 100 | 7.7 | 2.7 | 2.1 |
| 22 | M | 33 | AML | MAB / TreoFlu | pos. | pos. | Yes | No | Yes (83000) | Yes | 48 | 100 | N.A. | 1.6 | 0.8 |
| 23 | M | 59 | MDS | RIC / FBM | pos. | pos. | No | Yes (16000) | Yes (27000) | No | 30 | 99.3 | 3.4 | 1.6 | 1.3 |

Gender, age, disease type, conditioning regime, virus status (CMV and EBV) and clinical data (incidence of aGvHD, CMV and EBV reactivation, relapse and chimerism) of the corresponding stem cell recipients. In addition cell numbers for $CD4^+$ and $CD8^+$ T cells in the peripheral blood of donors and $CD34^+$ cells counts in the transplant are shown.

Threshold for antiviral treatment for CMV reactivation (Ganciclovir or Foscarnet) was >5000 copies/ml PB, for EBV reactivation (Rituximab) was >20000 copies/ml PB.

**ALL** indicates acute lymphatic leukemia; **AML**: acute myeloid leukemia; **CLL**: chronic lymphatic leukemia; **CML**: chronic myeloid leukemia; **MDS**: myelodysplastic syndrome; **MM**: multiple myeloma; **OMF**: osteomyelofibrosis; **MAB**: myeloablative conditioning; **RIC**: reduced intensity conditioning; **FBM**: Fludarabin 30mg/sqm d-9,-8,-7,-6,-5 + BCNU 150mg/sqm d-7,-6 + Melphalan 110mg/sqm d-4 + Thymoglobulin 2.5mg/kg (SIB), 7.5mg/kg (MUD); **TreoFlu**: Treosulfan 12g/sqm d-6,-5,-4 + Fludarabin 30mg/sqm d-6,-5,-4,-3,-2 + Thymoglobulin 2.5mg/kg (SIB), 7.5mg/kg (MUD); **BusFlu**: Busulfane i.v. 6.4mg/kg total dose + Fludarabin 30mg/sqm d-6,-5,-4 + Thymoglobulin 2.5mg/kg (SIB), 7.5mg/kg (MUD); **TTFluCy**: Thiotepa 10mg/kg d-6 + Fludarabin 30mg/kg d-4,-3 + Cyclophosphamide 30mg/kg d-4,-3 + Thymoglobulin 2.5mg/kg (SIB), 7.5mg/kg (MUD); **N.A.**: not available; **BW**: body weight; **PB**: peripheral blood

**Table S2.** Risk factors for aGvHD incidence.

| Recipient follow up | | acute GvHD | | | |
|---|---|---|---|---|---|
| | | Yes | No | total | p-value x2-test |
| Donor type | MUD | 12 | 3 | 15 | |
| | SIB | 5 | 3 | 8 | 0.3627 |
| Donor age | <40 | 5 | 2 | 7 | |
| | >40 | 12 | 4 | 16 | 0.8576 |
| **Patient** | | | | | |
| Age | <40 | 3 | 1 | 4 | |
| | >40 | 14 | 5 | 19 | 0.9566 |
| **CMV** | Yes | 8 | 3 | 11 | |
| reactivation | No | 9 | 3 | 12 | 0.9013 |
| **EBV** | Yes | 5 | 2 | 7 | |
| reactivation | No | 12 | 4 | 16 | 0.8576 |

Contingency table of the incidence of aGvHD in recipients after aHSCT in correlation with clinical parameters as donor age and type as well as Patient age and virus reactivation.

aGvHD: acute Graft-versus-host Disease; $x^2$-test: Chi-squared test

**Table S3.** Primer sequences.

| TCRß V segment(s) | Primer sequence |
|---|---|
| **V2** | CCCTACACGACGCTCTTCCGATCTCAAATTTCACTCTGAAGATCCGGTCCACAA |
| **V3-1** | CCCTACACGACGCTCTTCCGATCTCTCACTTAAATCTTCACATCAATTCCCTGG |
| **V4-1** | CCCTACACGACGCTCTTCCGATCTTTAAACCTTCACCTACACGCCCTGC |
| **V4-2/3** | CCCTACACGACGCTCTTCCGATCTTTATTCCTTCACCTACACACCCTGC |
| **V5-1** | CCCTACACGACGCTCTTCCGATCTCTCTGAGATGAATGTGAGCACCTTG |
| **V5-3** | CCCTACACGACGCTCTTCCGATCTCTCTGAGATGAATGTGAGTGCCTTG |
| **V5-4/5/6/7/8** | CCCTACACGACGCTCTTCCGATCTCTCTGAGCTGAATGTGAACGCCTTG |
| **V6-1** | CCCTACACGACGCTCTTCCGATCTCGCTCAGGCTGGAGTCGGCTG |
| **V6-2/3** | CCCTACACGACGCTCTTCCGATCTCTGGGGTTGGAGTCGGCTG |
| **V6-4** | CCCTACACGACGCTCTTCCGATCTCCTCACGTTGGCGTCTGCTG |
| **V6-5** | CCCTACACGACGCTCTTCCGATCTCTCAGGCTGCTGTCGGCTG |
| **V6-6** | CCCTACACGACGCTCTTCCGATCTGCTCAGGCTGGAGTTGGCTG |
| **V6-7** | CCCTACACGACGCTCTTCCGATCTCCCTCAAGCTGGAGTCAGCTG |
| **V6-8** | CCCTACACGACGCTCTTCCGATCTACTCAGGCTGGTGTCGGCTG |
| **V6-9** | CCCTACACGACGCTCTTCCGATCTGCTCAGGCTGGAGTCAGCTG |
| **V7-1** | CCCTACACGACGCTCTTCCGATCTCACTCTGAAGTTCCAGCGCACAC |
| **V7-2** | CCCTACACGACGCTCTTCCGATCTACTCTGACGATCCAGCGCACAC |
| **V7-3** | CCCTACACGACGCTCTTCCGATCTTCTACTCTGAAGATCCAGCGCACAG |
| **V7-4** | CCCTACACGACGCTCTTCCGATCTCACTCTGAAGATCCAGCGCACAG |
| **V7-6** | CCCTACACGACGCTCTTCCGATCTACTCTGACGATCCAGCGCACAG |
| **V7-7** | CCCTACACGACGCTCTTCCGATCTCACTCTGACGATTCAGCGCACAG |
| **V7-8** | CCCTACACGACGCTCTTCCGATCTCACTCTGAAGATCCAGCGCACAC |
| **V7-9** | CCCTACACGACGCTCTTCCGATCTACCTTGGAGATCCAGCGCACAG |
| **V9** | CCCTACACGACGCTCTTCCGATCTCACTCTGAACTAAACCTGAGCTCTCTG |
| **V10-1** | CCCTACACGACGCTCTTCCGATCTCCCTCACTCTGGAGTCTGCTG |
| **V10-2** | CCCTACACGACGCTCTTCCGATCTCCCCTCACTCTGGAGTCAGCTA |
| **V10-3** | CCCTACACGACGCTCTTCCGATCTCTCCTCACTCTGGAGTCCGCTA |
| **V11-1/3** | CCCTACACGACGCTCTTCCGATCTCACTCTCAAGATCCAGCCTGCAG |
| **V11-2** | CCCTACACGACGCTCTTCCGATCTTCCACTCTCAAGATCCAGCCTGCAA |
| **V12-3/4/5** | CCCTACACGACGCTCTTCCGATCTCACTCTGAAGATCCAGCCCTCAG |
| **V13** | CCCTACACGACGCTCTTCCGATCTATTCTGAACTGAACATGAGCTCCTTGG |
| **V14** | CCCTACACGACGCTCTTCCGATCTTACTCTGAAGGTGCAGCCTGCAG |
| **V15** | CCCTACACGACGCTCTTCCGATCTATAACTTCCAATCCAGGAGGCCGAACA |
| **V16** | CCCTACACGACGCTCTTCCGATCTTGTAGCCTTGAGATCCAGGCTACGA |
| **V17** | CCCTACACGACGCTCTTCCGATCTTTCCACGCTGAAGATCCATCCCG |
| **V18** | CCCTACACGACGCTCTTCCGATCTCATCCTGAGGATCCAGCAGGTAG |
| **V19** | CCCTACACGACGCTCTTCCGATCTCTCTCACTGTGACATCGGCCC |
| **V20-1** | CCCTACACGACGCTCTTCCGATCTTTGTCCACTCTGACAGTGACCAGTG |
| **V23-1** | CCCTACACGACGCTCTTCCGATCTAGCCTGGCAATCCTGTCCTCAG |

| | |
|---|---|
| **V24-1** | CCCTACACGACGCTCTTCCGATCTTCCCTGTCCCTAGAGTCTGCCAT |
| **V25-1** | CCCTACACGACGCTCTTCCGATCTCCTGACCCTGGAGTCTGCCA |
| **V27** | CCCTACACGACGCTCTTCCGATCTCCTGATCCTGGAGTCGCCCA |
| **V28** | CCCTACACGACGCTCTTCCGATCTTCCCTGATTCTGGAGTCCGCCA |
| **V29-1** | CCCTACACGACGCTCTTCCGATCTTAACATTCTCAACTCTGACTGTGAGCAACA |
| **V30** | CCCTACACGACGCTCTTCCGATCTGGCAGTTCATCCTGAGTTCTAAGAAGC |
| **TCRß J segment** | **Primer sequence** |
| **J1-1** | TTCAGACGTGTGCTCTTCCGATCTCTTACCTACAACTGTGAGTCTGGTGCC |
| **J1-2** | TTCAGACGTGTGCTCTTCCGATCTCTTACCTACAACGGTTAACCTGGTCCCCG |
| **J1-3** | TTCAGACGTGTGCTCTTCCGATCTCTCACCTACAACAGTGAGCCAACTTCCCT |
| **J1-4** | TTCAGACGTGTGCTCTTCCGATCTACCCAAGACAGAGAGCTGGGTTCCACT |
| **J1-5** | TTCAGACGTGTGCTCTTCCGATCTCTTACCTAGGATGGAGAGTCGAGTCC |
| **J1-6** | TTCAGACGTGTGCTCTTCCGATCTACCTGTCACAGTGAGCCTGGTCCCGT |
| **J2-1** | TTCAGACGTGTGCTCTTCCGATCTTACCTAGCACGGTGAGCCGTGTCCC |
| **J2-2** | TTCAGACGTGTGCTCTTCCGATCTCTTACCCAGTACGGTCAGCCTAGAGC |
| **J2-3** | TTCAGACGTGTGCTCTTCCGATCTGAGCACTGTCAGCCGGGTGCCTGG |
| **J2-4** | TTCAGACGTGTGCTCTTCCGATCTCAGCACTCAGAGCCGGGTCCCGGC |
| **J2-5** | TTCAGACGTGTGCTCTTCCGATCTACCGAGCACCAGGAGCCGCGTGC |
| **J2-6** | TTCAGACGTGTGCTCTTCCGATCTAGCACGGTCAGCCTGCTGCCGGC |
| **J2-7** | TTCAGACGTGTGCTCTTCCGATCTGTGACCGTGAGCCTGGTGCCCGG |
| **Adapter primer** | **Primer sequence** |
| **FW** | AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTC |
| **REV1** | CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGC |
| **REV2** | CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGC |
| **REV3** | CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTGC |
| **REV4** | CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGACGTGTGC |
| **REV5** | CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGC |
| **REV6** | CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGC |
| **REV7** | CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGTGC |
| **REV8** | CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGACGTGTGC |
| **REV9** | CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGC |
| **REV10** | CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGC |