Supplementary Materials for

Origin and evolution of the T-cell repertoire after posttransplantation cyclophosphamide

Christopher G. Kanakry,[1*] David G. Coffey,[2,4*] Andrea M.H. Towlerton,[2*] Ante Vulic,[1*] Barry E. Storer,[2] Jeffrey Chou,[2,4] Cecilia C.S. Yeung,[2,5] Christopher D. Gocke,[1] Harlan S. Robins,[3] Paul V. O'Donnell,[2,4] Leo Luznik,[1†] and Edus H. Warren[2,4†]

[1]Sidney Kimmel Comprehensive Cancer Center, The Johns Hopkins University School of Medicine, Baltimore, MD, U.S.A.; [2]Clinical Research and [3]Public Health Sciences Divisions, Fred Hutchinson Cancer Research Center, Seattle, WA, U.S.A.; Departments of [4]Medicine and [5]Pathology, University of Washington, Seattle, WA, U.S.A.

This PDF file includes:

Supplementary Methods.
Table S1. Clinical characteristics and *TRB* sequencing assessments performed for patients treated with either PTCy or CNI/MTX.
Table S2. Detailed characteristics of PTCy-treated patients on whose peripheral blood *TRB* sequencing was performed.
Table S3. Detailed characteristics of CNI/MTX-treated patients on whose peripheral blood *TRB* sequencing was performed.
Table S4. Detection of CMV-associated public *TRB* sequences in CMV tetramer$^+$ sorted blood.
Table S5. EBV-associated public *TRB* sequences in donor or recipient peripheral blood samples.
Table S6. Parameters of all unsorted peripheral blood samples assessed by *TRB* sequencing in these studies.
Table S7. Parameters of all flow cytometrically sorted PB or BM samples assessed by *TRB* sequencing in these studies.
Table S8. Parameters of all formalin-fixed and paraffin-embedded (FFPE) samples assessed by *TRB* sequencing in these studies.
Table S9. Parameters of all samples assessed by *TRA* sequencing in these studies.
Table S10. Parameters of all samples assessed by *IGH* sequencing in these studies.
Figure S1. Slow posttransplant recovery of CD4$^+$ recent thymic emigrants.
Figure S2. Recipient CMV seropositivity is associated with rapid CD3$^+$CD8$^+$ T-cell recovery in patients without detectable CMV reactivation.
Figure S3. *TRA* and *TRB* clonality are highly correlated.
Figure S4. Lorenz curves for *TRB* repertoires in blood samples from recipient 002-037 and his donor.
Figure S5. No correlation between donor or recipient age and *TRB* repertoire clonality.
Figure S6. Clonality of blood CD8$^+$ *TRB* repertoires is higher than that of blood CD4$^+$ repertoires.
Figure S7. *TRB* clonality is higher among memory T cells compared with naïve T cells and among memory CD8$^+$ T cells compared with memory CD4$^+$ T cells.
Figure S8. The Bhattacharyya coefficient and similarity score are highly correlated.
Figure S9. Clonality assessment of the *TRB* repertoire is robust to sequencing depth and reproducible.
Figure S10. Global similarity and sequence sharing between donor peripheral blood and bone marrow *TRB* repertoires.
Figure S11. CMV-associated *TRB* sequences are detectable in the blood and gastrointestinal tract.
Figure S12. Public EBV-specific *TRB* sequences are detectable.

Supplementary Methods

Antibodies used for flow cytometry

For flow cytometric immunophenotyping, monoclonal antibodies used included: Alexa Fluor (AF)-700 anti-CD3 (BioLegend, San Diego, CA), APC/Cy-7 anti-CD45RA (BioLegend), Brilliant Violet (BV)-650 anti-CD19 (BioLegend), BV-605 anti-CD4 (BD), BV-605 anti-CD38 (BioLegend), PE anti-CD31 (BD), FITC anti-CD56 (BD), PerCP-Cy5.5 anti-CD8 (BD), V450 anti-CCR7 (BD), PE-Cy7 anti-CD24 (BD), FITC anti-IgD (BD), and APC anti-CD27 (BD). Additional monoclonal antibodies used for the fluorescence-activated cell sorting (FACS) for sequencing studies included: PE anti-CCR7 (eBioscience, Inc., San Diego, CA), FITC anti-CD27 (eBioscience), PE-Cy7 anti-CD28 (eBioscience), APC anti-CD95 (eBioscience), and BV650 anti-CD45RO (Biolegend).

Genomic DNA extraction

Genomic DNA was isolated directly from fresh or cryopreserved/thawed unfractionated peripheral blood mononuclear cells or bone marrow mononuclear cells and from T-cell subsets sorted from peripheral blood or bone marrow using the DNAeasy Blood and Tissue Kit (Qiagen, Inc., Valencia, CA). For graft-versus-host disease target organ biopsies, genomic DNA was extracted from formalin-fixed, paraffin-embedded (FFPE) tissue sections using the Gentra Puregene Tissue Kit (Qiagen, Inc., Valencia, CA), with the primary modifications that Hemo-De was substituted for Xylene during washing, and the cell lysates were incubated on ice for 5 minutes after the addition of protein precipitation solution and before centrifugation.

Table S1. Clinical characteristics and *TRB* sequencing assessments performed for patients treated with either PTCy or CNI/MTX.

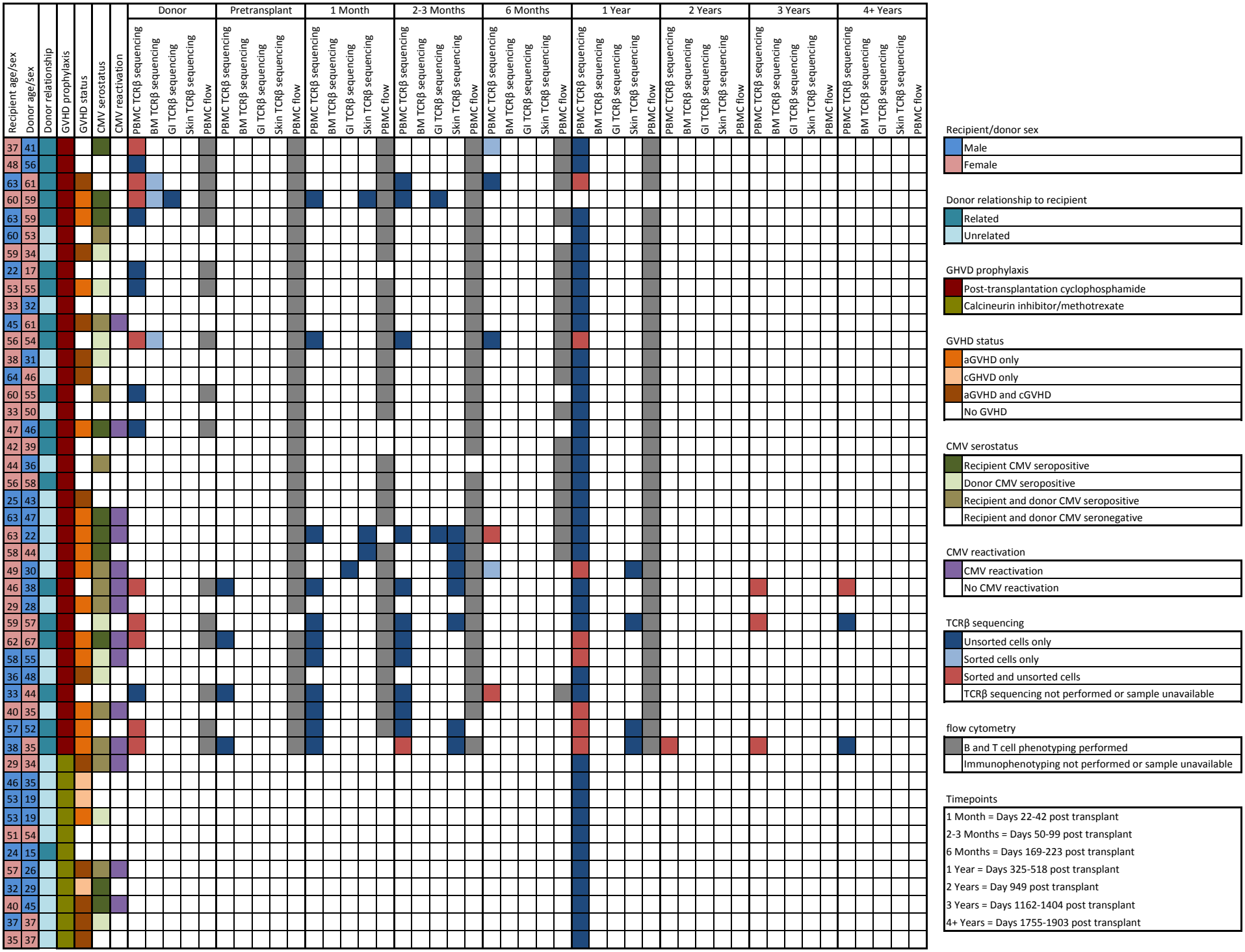

Table S2. Detailed characteristics of PTCy-treated patients on whose peripheral blood *TRB* sequencing was performed.

Patient No.	Age	Sex	Disease	Donor Type	Donor Age	Donor Sex	Donor PB Available/ Sequenced	Patient CMV Serology	Donor CMV Serology	PB CMV Reactivation Post-Transplant	Maximal aGVHD Grade	Chronic GVHD	Steroids Used for GVHD Treatment	NSSI Used for GVHD Treatment	Outcome[a]
001-001	37	F	AML	MRD	41	M	Yes	+	-	No	0	No	N/A	N/A	CR/Alive
001-005	48	F	AML	MRD	56	M	Yes	-	-	No	0	No	N/A	N/A	CR/Alive
001-006	63	M	AML	MRD	61	F	Yes[c]	-	-	No	II	Yes	Yes	Yes	CR/Alive
001-007[b]	60	F	AML	MRD	59	F	Yes[c]	+	-	No	III/IV	No	Yes	No	CR/Died
001-009	63	M	MDS	MRD	59	F	Yes	+	-	No	II	No	Yes	No	Relapsed/Alive
001-010	60	M	AML	MUD	53	F	No	+	+	No	0	No	N/A	N/A	Relapsed/Alive
001-012	59	F	MDS	MUD	34	F	No	-	+	No	III/IV	Yes	Yes	Yes	CR/Alive
001-015	22	M	AML	MRD	17	F	Yes	-	-	No	0	No	N/A	N/A	CR/Alive
001-016	53	F	AML	MRD	55	F	Yes	-	+	No	II	No	Yes	Yes	CR/Alive
001-017	33	F	AML	MUD	32	M	No	-	-	No	0	No	N/A	N/A	CR/Alive
001-022	45	M	ALL	MRD	61	F	No	+	+	Yes	II	Yes	Yes	Yes	CR/Alive
001-023	56	F	ALL	MRD	54	F	Yes[c]	-	+	No	0	No	N/A	N/A	Relapsed/Died
001-024	38	F	ALL	MUD	31	M	No	-	+	No	III/IV	Yes	Yes	Yes	CR/Alive
001-025	64	M	AML	MUD	46	F	No	-	-	No	II	Yes	Yes	Yes	CR/Alive
001-027	60	F	AML	MRD	55	F	Yes	+	+	No	0	No	N/A	N/A	CR/Alive
001-034	33	F	ALL	MUD	50	F	No	-	-	No	0	No	N/A	N/A	CR/Alive
001-035	47	F	AML	MRD	46	M	Yes	+	-	Yes	II	No	Yes	Yes	CR/Alive
001-037	42	F	ALL	MRD	39	F	No	-	-	No	0	No	N/A	N/A	CR/Alive
001-038	44	F	ALL	MUD	36	M	No	+	+	No	0	No	N/A	N/A	CR/Alive
001-040	56	F	ALL	MRD	58	F	No	-	-	No	0	No	N/A	N/A	Relapsed/Alive
001-041	25	M	ALL	MUD	43	M	No	-	-	No	III/IV	Yes	Yes	Yes	CR/Alive
001-044	63	M	AML	MUD	47	M	No	+	-	Yes	III/IV	No	Yes	Yes[d]	CR/Alive
002-001	63	F	AML	MUD	22	M	No	+	-	Yes	III/IV	No	Yes	Yes	Relapsed/Died
002-005	58	F	AML	MUD	44	F	No	+	-	No	II	No	Yes	Yes	CR/Alive
002-010	49	F	AML	MUD	30	M	No	+	+	Yes	II	No	No	Yes	Relapsed/Died
002-011	46	F	AML	MRD	38	M	Yes	+	+	Yes	0	No	N/A	N/A	CR/Alive
002-013	29	F	AML	MUD	28	M	No	+	+	Yes	II	No	Yes	No	CR/Alive
002-014	59	F	ALL	MRD	57	F	Yes	-	+	No	0	No	N/A	N/A	CR/Alive
002-016	62	F	AML	MRD	67	F	Yes	+	-	Yes	II	No	Yes	Yes	CR/Alive
002-017	58	M	AML	MUD	55	M	No	-	+	Yes	II	No	Yes	Yes	CR/Alive
002-018	36	M	AML	MUD	48	M	No	-	+	No	II	Yes	Yes	Yes	CR/Alive
002-019	33	M	AML	MRD	44	F	Yes	-	-	No	0	No	N/A	N/A	CR/Alive
002-022	40	F	AML	MUD	35	F	No	+	+	Yes	II	No	Yes	Yes	CR/Alive
002-031	57	M	AML	MRD	52	M	Yes	-	-	No	I	No	Yes	Yes	CR/Alive
002-037	38	M	AML	MRD	35	F	Yes	+	+	Yes	II	No	Yes	Yes	CR/Alive

[a]Follow-up of outcomes is through November 30, 2013.

[b]This patient died at day 64 of infection related to acute GVHD. Thus, a one-year sample was not available. However, donor PB and BM and patient 1 month and 2 month samples were sequenced.

[c]Donor bone marrow for these individuals from the time of allograft harvesting was also sequenced.

[d]Patient's only NSSI was mycophenolate mofetil versus placebo on clinical study. Whether the patient received the study drug or placebo has not been unblinded.

Abbreviations: PTCy, posttransplantation cyclophosphamide; No, number; PB, peripheral blood; CMV, cytomegalovirus; GVHD, graft-versus-host disease; aGVHD, acute GVHD; NSSI, non-steroidal systemic immunosuppressant(s); F, female; M, male; AML, acute myeloid leukemia; MDS, myelodysplastic syndrome; CML, chronic myelogenous leukemia; ALL, acute lymphoblastic leukemia; MRD, HLA-matched-related donor; MUD, HLA-matched-unrelated donor; N/A, non-applicable; CR, complete remission; NRM, non-relapse mortality.

Table S3. Detailed characteristics of CNI/MTX-treated patients on whose peripheral blood *TRB* sequencing was performed.

Patient No.	Age	Sex	Disease	Conditioning	Donor Type	Donor Age	Donor Sex	Patient CMV Serology	Donor CMV Serology	PB CMV Reactivation Post-Transplant	Maximal aGVHD Grade	Chronic GVHD	Steroids Used for GVHD Treatment	NSSI Used for GVHD Treatment	Outcome[a]
003-001	29	F	ALL	Cy/TBI	MUD	34	F	+	+	Yes	II	Yes	Yes	Yes	CR/Alive
003-002	45	M	AML	Bu/Cy	MUD	35	M	-	-	No	0	Yes	Yes	Yes	CR/Alive
003-003	53	M	AML	Treo/Flu/TBI	MUD	19	M	-	-	No	0	Yes	Yes	Yes	CR/Alive
003-004	31	M	ALL	Cy/TBI	MUD	35	M	-	+	No	III	No	Yes	Yes	Relapsed/Died
003-005	50	F	AML	Treo/Flu/TBI	MUD	54	F	-	-	No	0	No	No	No	CR/Alive
003-006	23	M	MDS	Bu/Cy	MRD	15	M	-	-	No	0	No	No	No	CR/Alive
003-007	57	F	ALL	Cy/TBI	MUD	26	M	+	+	Yes	II	Yes	Yes	Yes	CR/Died
003-008	32	M	ALL	Cy/TBI	MUD	29	M	+	-	No	0	Yes	No	No	CR/Alive
003-009	39	F	CML	Bu/Cy	MUD	45	M	+	-	Yes	II	Yes	Yes	Yes	CR/Alive
003-010	37	M	ALL	Cy/TBI	MUD	37	F	-	+	No	III	Yes	Yes	Yes	CR/Alive
003-011	34	F	CML	Bu/Cy	MUD	37	F	-	-	No	II	Yes	Yes	Yes	CR/Alive

[a]Follow-up of outcomes is through March 18, 2015.

Abbreviations: CNI, calcineurin-inhibitor; MTX, methotrexate; No., number; PB, peripheral blood; CMV, cytomegalovirus; GVHD, graft-versus-host disease; aGVHD, acute GVHD; NSSI, non-steroidal systemic immunosuppressant(s); F, female; M, male; AML, acute myeloid leukemia; MDS, myelodysplastic syndrome; CML, chronic myelogenous leukemia; ALL, acute lymphoblastic leukemia; Cy, cyclophosphamide; TBI, total body irradiation; Bu, busulfan; Treo, Treosulfan; Flu, fludarabine; MRD, HLA-matched-related donor; MUD, HLA-matched-unrelated donor; N/A, non-applicable; CR, complete remission; NRM, non-relapse mortality.

Table S4. Detection of CMV-associated public *TRB* sequences in CMV tetramer^{+} sorted blood.

Recipient ID	Donor/Recipient CMV serostatus	CMV Reactivation	Day of blood collection	CMV tetramer used for sorting	Donor/Recipient HLA A and B alleles				Public CMV TCR			Frequency (%) of public CMV TCR detected in CMV tetramer+ sorted blood
									CDR3β sequence	Epitope	HLA restriction	
002-010	+/+	Yes	369	CMV HLA-A2 NLVPMVATV	HLA-A3	HLA-A2	HLA-B49	HLA-B7	CASSLAPGATNEKLFF	CMV pp65495-503	HLA-A2	10.42
002-022	+/+	Yes	365	CMV HLA-A2 NLVPMVATV	HLA-A2	HLA-A24	HLA-B40	HLA-B35	CASSSANYGYTF	CMV pp65495-503	HLA-A2	18.49
									CASSLEGYTEAFF			0.01
002-037	+/+	Yes	83	CMV HLA-B7 RPHERNGFTVL	HLA-A1	HLA-A2	HLA-B7	HLA-B35	CASSPSRNTEAFF	CMV pp65265-275	HLA-B7	2.79
									CASSPARNTEAFF			1.43
002-037	+/+	Yes	369	CMV HLA-B7 RPHERNGFTVL	HLA-A1	HLA-A2	HLA-B7	HLA-B35	CASSPSRNTEAFF	CMV pp65265-275	HLA-B7	0.88
									CASSPARNTEAFF	CMV pp65265-275	HLA-B7	1.14

Table S5. EBV-associated public *TRB* sequences detected in posttransplant peripheral blood samples from recipient 002-014.

CDR3β Sequence	Phenotype	HLA	EBV Epitope	Reference (Pubmed ID)
CAISTGDSNQPQHF	$CD8^+$	HLA-B*3501	BZLF1-EPLPQGQLTAY	16148129
CASGTGDSNQPQHF	$CD8^+$	HLA-B*3501	BZLF1-EPLPQGQLTAY	16148129
CASSARSGELFF	Not reported	HLA-B*3501	EBNA1-HPVGEADYFEY	17082594
CASSARTGELFF	Not reported	HLA-B*3501	EBNA1-HPVGEADYFEY	17082594
CASSDSSLTEAFF	Not reported	HLA-B*35	EBNA3A-YPLHEQHGM	23267020
CASSLAGANVLTF	Not reported	HLA-B8	BZLF1-RAKFKQLLQ	9418135
CASSLGQAYEQYF	$CD8^+$	HLA-B*08:01	EBNA-3A325-333	7964506
CASSPGNYEQYF	Not reported	HLA-B8	EBNA-3-FLRGRAYGL	9788968
CASSQGVGENTEAFF	Not reported	HLA-B*08	BZLF1-RAKFKQLL	23267020
CASSSGSYEQYF	Not reported	HLA-B8	EBNA-3-FLRGRAYGL	9788968
CASSSRGNEQFF	Not reported	HLA-B8	EBNA-3-FLRGRAYGL	9788968
CASSTGDSNQPQHF	$CD8^+$	HLA-B*3501	BZLF1-EPLPQGQLTAY	16148129
CASSVRTGELFF	Not reported	HLA-B*3501	EBNA1-HPVGEADYFEY	17082594
CATGTGDSNQPQHF	$CD8^+$	HLA-B*3501	BZLF1-EPLPQGQLTAY	16148129
CATSTGDSNQPQHF	Not reported	HLA-B*3501	BZLF1-EPLPQGQLTAY	16148129
CSVGSGEGYEQYF	Not reported	HLA-B8	BZLF1-RAKFKQLLQ	9418135

Table S6. Parameters of all unsorted peripheral blood samples assessed by *TRB* sequencing in these studies (page 1 of 2).

Recipient ID	Time point	Tissue source	Number of unique nucleotide sequences per sample	Number of unique, productive amino acid sequences per sample	Estimated number of genomes per sample	Number of reads per sample
001-001	1 Year	PBMC	34,322	27,835	78,461	3,194,921
001-001	Donor	PBMC	93,482	78,277	225,941	5,927,888
001-005	1 Year	PBMC	22,582	17,796	90,478	1,967,236
001-005	Donor	PBMC	52,993	41,740	130,723	3,110,368
001-006	2-3 Months	PBMC	89,207	73,785	106,773	1,607,947
001-006	6 Months	PBMC	6,870	5,537	12,043	266,864
001-006	1 Year	PBMC	6,339	4,634	20,919	1,716,864
001-006	Donor	PBMC	49,360	39,932	153,111	2,208,247
001-007	1 Month	PBMC	6,010	4,914	22,536	620,142
001-007	2-3 Months	PBMC	4,777	3,847	14,251	570,680
001-007	Donor	PBMC	56,559	47,150	72,819	1,149,806
001-009	1 Year	PBMC	12,152	9,217	96,876	1,963,682
001-009	Donor	PBMC	52,056	40,812	133,897	4,818,935
001-010	1 Year	PBMC	11,218	8,609	44,400	951,536
001-012	1 Year	PBMC	7,341	5,725	11,647	300,288
001-015	1 Year	PBMC	51,242	40,464	120,435	2,567,505
001-015	Donor	PBMC	74,092	59,341	190,918	1,292,902
001-016	1 Year	PBMC	9,053	7,117	62,589	1,830,388
001-016	Donor	PBMC	57,168	46,983	160,371	1,046,106
001-017	1 Year	PBMC	22,093	18,652	26,626	663,853
001-022	1 Year	PBMC	9,675	7,373	52,077	1,269,333
001-023	1 Month	PBMC	5,540	4,391	30,034	1,505,287
001-023	2-3 Months	PBMC	6,218	4,973	23,089	1,143,159
001-023	6 Months	PBMC	8,536	6,808	16,035	961,724
001-023	1 Year	PBMC	13,706	11,012	27,706	1,266,072
001-023	Donor	PBMC	91,923	75,883	118,437	1,443,727
001-024	1 Year	PBMC	11,815	9,326	89,741	1,437,164
001-025	1 Year	PBMC	4,027	3,024	10,813	547,462
001-027	1 Year	PBMC	11,937	9,803	144,487	1,487,423
001-027	Donor	PBMC	78,462	64,594	230,401	3,489,581
001-034	1 Year	PBMC	44,511	35,976	72,565	1,114,156
001-035	1 Year	PBMC	25,693	20,781	178,721	1,785,112
001-035	Donor	PBMC	90,289	75,108	251,103	2,525,698
001-037	1 Year	PBMC	37,545	30,765	52,686	713,876
001-038	1 Year	PBMC	17,696	14,345	40,569	1,156,872
001-040	1 Year	PBMC	21,689	17,405	36,071	903,749
001-041	1 Year	PBMC	15,794	12,962	53,269	1,740,556
001-044	1 Year	PBMC	10,011	7,884	57,437	817,428
002-001	1 Month	PBMC	5,112	4,226	NA	1,168,269
002-001	2-3 Months	PBMC	14,281	11,789	NA	580,280
002-001	6 Months	PBMC	2,989	2,477	22,618	726,742
002-001	1 Year	PBMC	4,332	3,621	NA	769,689
002-005	1 Year	PBMC	15,410	12,150	35,842	671,731
002-010	1 Year	PBMC	27,600	21,433	231,554	2,724,110
002-011	1 Month	PBMC	3,514	2,785	NA	641,019
002-011	2-3 Months	PBMC	6,016	4,807	NA	1,160,372
002-011	1 Year	PBMC	19,376	15,633	NA	1,325,474
002-011	3 Years	PBMC	28,315	23,040	125,061	1,821,146
002-011	4+ Years	PBMC	20,689	16,530	39,967	39,967
002-011	Donor	PBMC	69,972	57,072	NA	1,583,656
002-011	Pretransplant	PBMC	66,194	54,627	101,550	2,127,214
002-013	1 Year	PBMC	8,354	6,672	34,668	849,237
002-014	1 Month	PBMC	4,880	4,087	NA	621,579

Table S6. Parameters of all unsorted peripheral blood samples assessed by *TRB* sequencing in these studies (page 2 of 2).

Recipient ID	Time point	Tissue source	Number of unique nucleotide sequences per sample	Number of unique, productive amino acid sequences per sample	Estimated number of genomes per sample	Number of reads per sample
002-014	2-3 Months	PBMC	5,287	4,495	NA	449,199
002-014	1 Year	PBMC	14,561	12,064	NA	1,001,596
002-014	3 Years	PBMC	13,724	11,190	17,359	1,551,053
002-014	5 Years	PBMC	20,659	17,413	35,936	NA
002-014	Donor	PBMC	62,434	53,288	NA	1,487,414
002-016	Pretransplant	PBMC	50,095	40,681	72,365	1,260,470
002-016	1 Month	PBMC	3,676	2,919	NA	427,965
002-016	2-3 Months	PBMC	4,121	3,232	NA	535,670
002-016	1 Year	PBMC	6,386	4,934	NA	1,013,709
002-016	Donor	PBMC	58,159	46,170	NA	1,101,415
002-017	1 Month	PBMC	8,620	6,895	25,319	2,118,919
002-017	2-3 Months	PBMC	9,231	7,438	19,088	2,189,115
002-017	1 Year	PBMC	14,550	11,913	116,390	1,620,091
002-018	1 Year	PBMC	14,095	11,375	37,761	895,105
002-018	1 Year (Repeat)	PBMC	14,550	11,899	NA	835,083
002-019	Pretransplant	PBMC	104,693	85,156	154,790	1,564,202
002-019	1 Month	PBMC	1,733	1,343	NA	316,850
002-019	2-3 Months	PBMC	2,347	1,854	NA	357,843
002-019	6 Months	PBMC	13,143	9,967	64,320	2,409,768
002-019	1 Year	PBMC	25,854	19,940	101,651	2,605,175
002-019	Donor	PBMC	24,939	19,556	NA	1,014,729
002-022	1 Month	PBMC	2,437	1,942	7,257	493,833
002-022	2-3 Months	PBMC	1,648	1,260	6,477	617,255
002-022	1 Year	PBMC	17,442	14,185	167,760	1,780,771
002-031	1 Month	PBMC	4,635	3,592	12,447	1,932,956
002-031	2-3 Months	PBMC	10,328	8,022	14,078	1,473,611
002-031	1 Year	PBMC	19,750	15,709	61,510	1,181,114
002-031	Donor	PBMC	78,611	64,218	204,459	2,009,271
002-037	Pretransplant	PBMC	5,768	4,614	23,542	1,347,901
002-037	1 Month	PBMC	920	767	NA	31,078
002-037	2-3 Months	PBMC	2,816	2,266	NA	484,664
002-037	10 Months	PBMC	4,177	3,255	11,869	963,151
002-037	1 Year	PBMC	6,515	5,216	NA	910,229
002-037	2 Years	PBMC	5,314	4,120	11,223	1,769,553
002-037	3 Years	PBMC	39,085	31,874	234,849	NA
002-037	4+ Years	PBMC	13,760	11,120	50,538	NA
002-037	Donor	PBMC	61,256	50,271	NA	1,233,683
002-037	Donor (Deep)	PBMC	109,400	87,390	139,094	NA
003-001	1 Year	PBMC	35,037	27,721	118,939	2,019,674
003-002	1 Year	PBMC	9,657	7,839	22,405	94,409
003-003	1 Year	PBMC	32,130	25,758	39,978	1,372,014
003-004	1 Year	PBMC	24,063	20,458	35,566	171,042
003-005	1 Year	PBMC	43,316	36,332	56,070	1,546,452
003-006	1 Year	PBMC	31,382	25,389	54,851	1,924,325
003-007	1 Year	PBMC	7,919	6,458	113,165	1,735,843
003-008	1 Year	PBMC	21,480	16,826	69,472	1,417,171
003-009	1 Year	PBMC	43,959	35,531	51,388	1,318,002
003-010	1 Year	PBMC	33,682	26,605	51,104	3,782,386
003-011	1 Year	PBMC	45,453	36,546	56,775	1,630,649
Total			**2,705,814**	**2,194,735**	**6,112,110**	**137,891,430**
Mean			**26,528**	**21,517**	**77,368**	**1,407,055**
Median			**14,556**	**11,989**	**54,851**	**1,267,703**
SD			**26,531**	**21,750**	**64,294**	**948,345**
Max			**109,400**	**87,390**	**251,103**	**5,927,888**
Min			**920**	**767**	**6,477**	**31,078**

Table S7. Parameters of all flow cytometrically sorted peripheral blood or bone marrow samples assessed by *TRB* sequencing in these studies (page 1 of 2).

Recipient ID	Time point	Tissue source	Immunophenotype	Number of unique nucleotide sequences per sample	Number of unique, productive amino acid sequences per sample	Estimated number of genomes per sample	Number of reads per sample
001-001	6 Months	PBMC	CD4	1,766	1,370	3,880	414,350
001-001	6 Months	PBMC	CD8	300	220	1,654	209,382
001-001	Donor	PBMC	CD4 Memory	12,133	10,062	12,313	488,940
001-001	Donor	PBMC	CD4 Naïve	39,959	34,088	40,215	875,000
001-001	Donor	PBMC	CD8 Memory	1,329	1,064	2,113	188,860
001-001	Donor	PBMC	CD8 Naïve	38,612	32,976	38,817	1,142,384
001-006	1 Year	PBMC	CD4	34,144	27,844	37,041	1,056,468
001-006	1 Year	PBMC	CD8	13,468	10,968	16,129	807,253
001-006	Donor	BM	CD4	11,505	9,183	12,194	771,791
001-006	Donor	BM	CD8	6,755	5,419	14,105	1,057,176
001-006	Donor	PBMC	CD4 Memory	19,316	15,538	22,209	917,519
001-006	Donor	PBMC	CD4 Naïve	20,255	16,410	20,404	992,613
001-006	Donor	PBMC	CD8 Memory	6,558	5,278	13,737	863,423
001-006	Donor	PBMC	CD8 Naïve	2,870	2,349	2,900	193,872
001-007	Donor	BM	CD4	1,023	784	1,120	76,776
001-007	Donor	BM	CD8	925	681	1,318	133,705
001-007	Donor	PBMC	CD4 Memory	28,978	24,020	44,360	729,271
001-007	Donor	PBMC	CD4 Naïve	8,070	6,673	8,379	355,481
001-007	Donor	PBMC	CD8 Memory	7,223	5,961	21,940	772,905
001-007	Donor	PBMC	CD8 Naïve	559	408	574	41,703
001-023	1 Year	PBMC	CD4	1,122	864	1,244	107,633
001-023	1 Year	PBMC	CD8	816	562	2,894	272,315
001-023	Donor	BM	CD4	3,900	3,088	4,208	341,939
001-023	Donor	BM	CD8	3,462	2,729	7,058	585,493
001-023	Donor	PBMC	CD4	14,783	11,983	17,357	966,116
001-023	Donor	PBMC	CD8	5,552	4,460	13,979	608,884
002-001	6 Months	PBMC	CD4	3,347	2,638	7,378	1,223,652
002-001	6 Months	PBMC	CD8	797	600	37,590	2,376,233
002-010	1 Year	PBMC	CD8 CMV tetramer A2 NLVPMVATV	56	30	528	NA
002-011	3 Years	PBMC	CD4	16,714	13,462	33,780	1,293,242
002-011	3 Years	PBMC	CD8	7,693	6,389	30,623	1,586,804
002-011	4+ Years	PBMC	CD4	74,729	60,813	134,378	134,378
002-011	4+ Years	PBMC	CD8	32,882	26,940	139,441	139,441
002-011	Donor	PBMC	CD4 Memory	36,997	30,003	56,191	769,683
002-011	Donor	PBMC	CD4 Naïve	64,753	53,509	71,503	647,931
002-011	Donor	PBMC	CD8 Memory	8,391	6,742	50,429	1,152,926
002-011	Donor	PBMC	CD8 Naïve	22,012	18,000	22,972	703,078
002-014	3 Years	PBMC	CD4	16,595	13,939	18,439	1,155,749
002-014	3 Years	PBMC	CD8	12,884	10,787	39,527	1,963,050
002-014	Donor	PBMC	CD4 Memory	13,586	11,389	30,593	640,323
002-014	Donor	PBMC	CD4 Naïve	30,943	26,418	32,799	914,162
002-014	Donor	PBMC	CD8 Memory	1,899	1,517	6,446	538,525
002-014	Donor	PBMC	CD8 Naïve	6,369	5,370	6,654	502,692
002-016	1 Year	PBMC	CD4	3,815	2,767	8,748	1,851,232
002-016	1 Year	PBMC	CD8	2,244	1,565	38,405	3,349,204
002-016	Donor	PBMC	CD4 Memory	45,834	36,237	72,784	1,108,108
002-016	Donor	PBMC	CD4 Naïve	65,262	51,659	72,954	1,118,504
002-016	Donor	PBMC	CD8 Memory	9,556	7,477	40,583	1,004,426
002-016	Donor	PBMC	CD8 Naïve	4,106	3,165	4,367	288,124
002-017	1 Year	PBMC	CD4	3,316	2,576	5,807	1,199,190
002-017	1 Year	PBMC	CD8	1,880	1,488	21,361	1,813,278
002-019	2-3 Months	PBMC	CD8	1,966	1,462	12,236	1,427,718
002-019	6 Months	PBMC	CD4	9,075	6,893	12,885	1,384,948

Table S7. Parameters of all flow cytometrically sorted peripheral blood or bone marrow samples assessed by *TRB* sequencing in these studies (page 2 of 2).

Recipient ID	Time point	Tissue source	Immunophenotype	Number of unique nucleotide sequences per sample	Number of unique, productive amino acid sequences per sample	Estimated number of genomes per sample	Number of reads per sample
002-019	6 Months	PBMC	CD8	4,245	3,282	24,771	1,475,658
002-022	1 Year	PBMC	CD8 CMV tetramer A2 NLVPMVATV	248	166	30,147	825,944
002-022	2-3 Months	PBMC	CD4	89	29	96	8,822
002-022	2-3 Months	PBMC	CD8	92	18	99	6,312
002-031	1 Year	PBMC	CD4	12,756	9,985	16,975	2,079,365
002-031	1 Year	PBMC	CD8	4,502	3,407	8,565	1,799,581
002-031	Donor	PBMC	CD4 Memory	27,754	22,850	32,790	839,322
002-031	Donor	PBMC	CD4 Naïve	46,069	37,751	46,686	1,084,937
002-031	Donor	PBMC	CD8 Memory	2,733	2,146	6,321	455,107
002-031	Donor	PBMC	CD8 Naïve	18,305	14,892	18,390	660,446
002-037	1 Year	PBMC	CD8 CMV tetramer B7 RPHERNGFTVL	414	281	1,794	52,480
002-037	2-3 Months	PBMC	CD8 CMV tetramer B7 RPHERNGFTVL	201	122	254	4,553
002-037	3 Years	PBMC	CD8 CMV tetramer A2 NLVPMVATV	40	25	53	NA
002-037	2 Years	PBMC	CD4	23,396	18,849	50,342	2,435,948
002-037	2 Years	PBMC	CD8	3,571	2,816	29,399	2,217,576
002-037	Donor	PBMC	CD8 Memory	6,833	5,242	34,946	1,143,208
002-037	Donor	PBMC	CD8 Naïve	13,766	11,223	13,881	318,355
Total				**948,098**	**771,901**	**1,687,052**	**60,665,467**
Mean				**13,544**	**11,027**	**24,101**	**892,139**
Median				**6,794**	**5,395**	**16,552**	**816,599**
SD				**16,844**	**13,789**	**26,969**	**679,697**
Max				**74,729**	**60,813**	**139,441**	**3,349,204**
Min				**40**	**18**	**53**	**4,553**

Table S8. Parameters of all formalin-fixed and paraffin-embedded (FFPE) samples assessed by *TRB* sequencing in these studies.

Recipient ID	Time point	Tissue source	Number of unique nucleotide sequences per sample	Number of unique, productive amino acid sequences per sample	Estimated number of genomes per sample	Number of reads per sample
001-007	1 Month	Skin	18	9	24	NA
001-007	2-3 Months	Antrum	25	9	28	NA
001-007	2-3 Months	Colon	25	8	27	NA
001-007	2-3 Months	Colon	19	7	22	NA
001-007	2-3 Months	Colon	23	9	24	NA
001-007	2-3 Months	Duodenum	32	14	35	NA
001-007	2-3 Months	Esophagus	24	3	26	NA
001-007	2-3 Months	GastricBody	58	40	62	NA
001-007	2-3 Months	Rectum	37	23	39	NA
001-007	Donor	Colon	503	407	569	NA
001-007	Donor	Colon	1,016	830	1,194	NA
001-007	Donor	Colon	1,618	1,333	2,222	NA
001-007	Donor	Ileum	392	314	435	NA
001-007	Donor	Rectum	1,471	1,231	2,079	NA
002-001	1 Month	Skin	77	25	102	8,553
002-001	2-3 Months	Antrum	123	62	139	8,162
002-001	2-3 Months	Antrum	80	18	84	2,138
002-001	2-3 Months	Duodenum	85	34	96	1,204
002-001	2-3 Months	Sigmoid	89	37	94	4,215
002-001	2-3 Months	Skin	81	20	108	7,523
002-005	1 Month	Skin	80	23	82	9,054
002-005	2-3 Months	Skin	71	14	80	7,960
002-010	1 Month	Antrum	147	79	152	16,606
002-010	1 Year	Skin	69	14	76	10,833
002-010	1 Year	Skin	97	26	110	12,851
002-010	2-3 Months	Skin	84	31	117	12,751
002-011	2-3 Months	Skin	76	14	86	11,345
002-014	1 Year	Skin	94	45	110	12,075
002-014	2-3 Months	Skin	156	96	196	13,351
002-031	1 Month	Skin	119	52	137	11,418
002-031	1 Year	Skin	84	18	92	10,362
002-031	2-3 Months	Skin	74	16	94	8,627
002-037	1 Year	Skin	320	236	507	62,418
002-037	2-3 Months	Antrum	119	50	134	13,267
002-037	2-3 Months	Antrum	223	147	256	18,267
002-037	2-3 Months	Duodenum	147	73	186	16,602
002-037	2-3 Months	Sigmoid	194	118	253	17,694
002-037	2-3 Months	Skin	72	25	87	8,902
Total			**8,022**	**5,510**	**10,164**	**306,178**
Mean			**211**	**145**	**267**	**12,757**
Median			**84**	**29**	**99**	**11,089**
SD			**365**	**311**	**498**	**11,457**
Max			**1,618**	**1,333**	**2,222**	**62,418**
Min			**18**	**3**	**22**	**1,204**

Table S9. Parameters of all samples assessed by *TRA* sequencing in these studies.

Recipient ID	Time point	Tissue source	Number of unique nucleotide sequences per sample	Number of unique, productive amino acid sequences per sample	Number of reads per sample
001-006	1 Year	PBMC	10,268	5,540	24,936
001-006	Donor	PBMC	89,741	49,636	163,469
002-001	1 Year	PBMC	6,221	3,551	68,660
002-001	6 Months	PBMC	5,757	3,250	20,467
002-001	6 Months	PBMC	1,076	544	128,021
002-014	1 Year	PBMC	34,196	18,837	84,347
002-014	3 Years	PBMC	32,985	17,814	50,253
002-014	3 Years	PBMC	32,766	15,302	130,749
002-031	Donor	PBMC	144,815	80,163	238,329
Total			**357,825**	**194,637**	**909,231**
Mean			**39,758**	**21,626**	**101,026**
Median			**32,766**	**15,302**	**84,347**
SD			**47,838**	**26,554**	**71,192**
Max			**144,815**	**80,163**	**238,329**
Min			**1,076**	**544**	**20,467**

Table S10. Parameters of all samples assessed by *IGH* sequencing in these studies.

Recipient ID	Time point	Tissue source	Number of unique nucleotide sequences per sample	Number of unique, productive amino acid sequences per sample	Number of reads per sample
001-001	1 Year	PBMC	97,046	80,791	1,430,028
001-001	Donor	PBMC	25,755	21,244	434,858
001-005	1 Year	PBMC	50,728	43,323	1,042,827
001-005	Donor	PBMC	14,313	12,059	506,471
001-006	1 Year	PBMC	51,301	42,838	948,520
001-006	Donor	PBMC	6,951	5,685	221,984
001-009	1 Year	PBMC	109,495	89,987	1,576,215
001-009	Donor	PBMC	25,334	20,817	439,606
001-012	1 Year	PBMC	28,253	23,992	449,850
001-015	1 Year	PBMC	73,498	61,226	1,537,039
001-015	Donor	PBMC	14,520	12,424	353,473
001-016	1 Year	PBMC	6,137	5,394	371,113
001-016	Donor	PBMC	27,066	22,783	532,364
001-022	1 Year	PBMC	13,399	11,790	341,532
001-024	1 Year	PBMC	24,432	21,354	950,547
001-025	1 Year	PBMC	59,596	48,485	1,490,098
001-027	1 Year	PBMC	18,322	15,004	645,977
001-027	Donor	PBMC	28,552	23,705	313,610
001-035	1 Year	PBMC	15,806	13,374	520,701
001-035	Donor	PBMC	10,408	8,772	100,281
001-041	1 Year	PBMC	16,100	13,983	772,988
002-001	1 Year	PBMC	5,248	4,535	107,834
002-005	1 Year	PBMC	41,641	35,417	977,274
002-010	1 Year	PBMC	51,203	42,409	1,869,794
002-011	1 Year	PBMC	48,492	40,987	575,255
002-011	3 Years	PBMC	51,838	43,018	1,148,404
002-011	Donor	PBMC	11,682	9,994	296,695
002-013	1 Year	PBMC	10,477	8,392	1,071,826
002-014	1 Year	PBMC	73,737	61,847	700,088
002-014	3 Years	PBMC	18,380	15,498	132,012
002-014	Donor	PBMC	26,441	21,900	308,599
002-016	1 Year	PBMC	13,184	11,274	273,789
002-016	Donor	PBMC	7,458	6,154	208,548
002-017	1 Year	PBMC	18,485	15,705	572,570
002-018	1 Year	PBMC	19,806	16,873	1,657,538
002-019	1 Year	PBMC	25,552	22,124	482,012
002-019	Donor	PBMC	7,624	6,392	149,040
002-022	1 Year	PBMC	58,575	48,876	899,132
002-031	1 Year	PBMC	97,547	79,505	2,306,976
002-031	Donor	PBMC	25,340	20,239	487,090
002-037	1 Year	PBMC	48,124	40,426	600,824
002-037	Donor	PBMC	17,820	14,733	338,203
003-002	1 Year	PBMC	44,775	37,503	1,228,143
003-003	1 Year	PBMC	82,446	68,166	1,006,688
003-004	1 Year	PBMC	61,103	52,417	1,090,991
003-005	1 Year	PBMC	66,744	56,487	734,270
003-006	1 Year	PBMC	44,458	37,911	654,015
003-007	1 Year	PBMC	4,702	3,933	283,501
003-009	1 Year	PBMC	62,626	51,741	447,277
003-010	1 Year	PBMC	507	407	77,339
Total			**1,763,027**	**1,473,893**	**35,665,809**
Mean			**35,261**	**29,478**	**713,316**
Median			**25,654**	**21,627**	**552,467**
SD			**27,193**	**22,498**	**509,473**
Max			**109,495**	**89,987**	**2,306,976**
Min			**507**	**407**	**77,339**

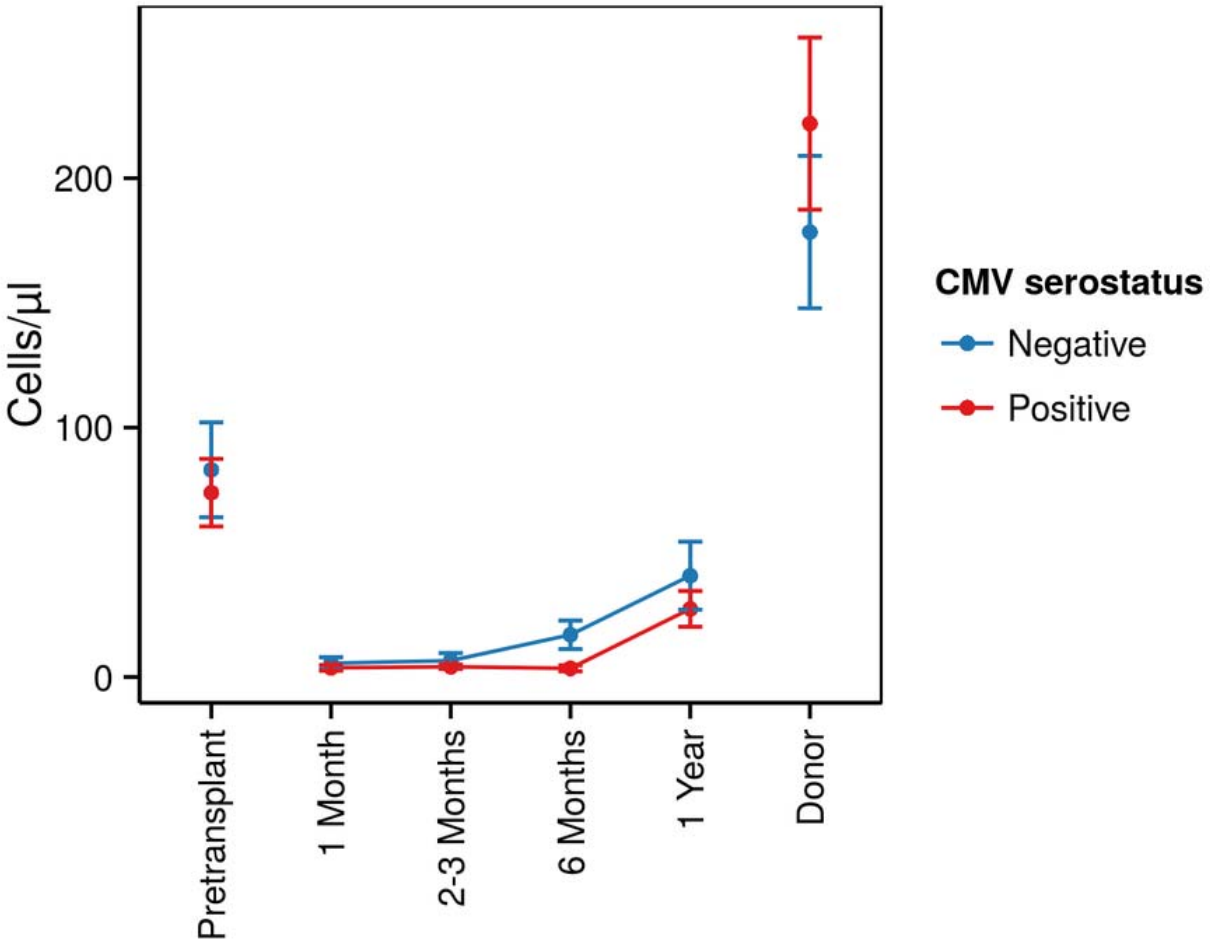

Figure S1. Slow posttransplant recovery of $CD4^+$ recent thymic emigrants. Numerical reconstitution of $CD3^+CD4^+CD45RA^+CD31^+$ cells is shown. Except for donor data, CMV serostatus refers to the recipient for all timepoints.

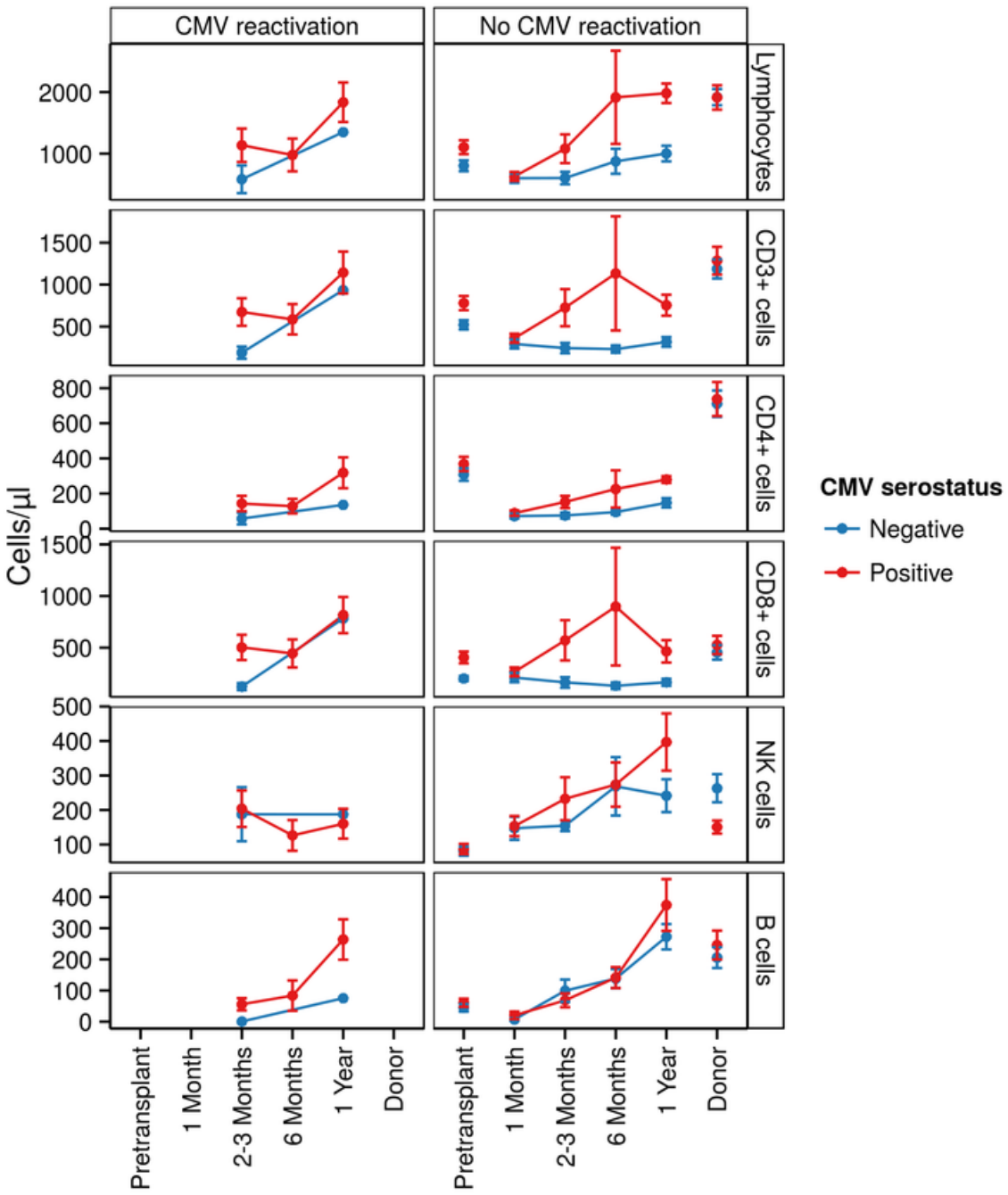

Figure S2. Recipient CMV seropositivity is associated with rapid $CD3^+CD8^+$ T-cell recovery in patients without detectable CMV reactivation. Although rapid $CD8^+$ T-cell numerical recovery is seen in patients who are CMV-seropositive or who had detectable CMV reactivation in the peripheral blood, recipient CMV seropositivity seemed to exert the dominant effect; patients without detectable CMV reactivation who were CMV-seropositive had more rapid $CD8^+$ and $CD4^+$ T-cell numerical reconstitution than did patients who were CMV-seronegative. Except for donor data, CMV serostatus refers to the recipient for all timepoints.

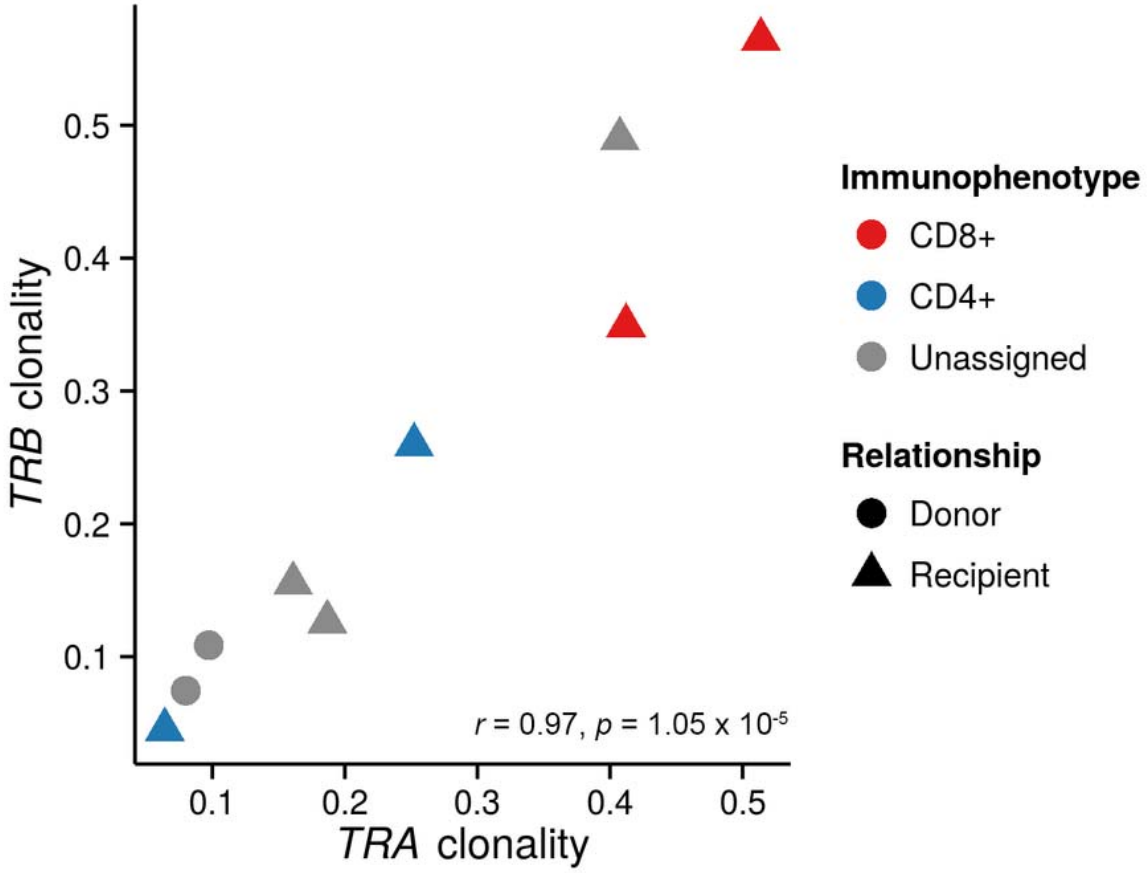

Figure S3. *TRA* and *TRB* clonality are highly correlated. *TRA* and *TRB* sequencing were performed on genomic DNA extracted from unsorted PBMC or from flow cytometrically sorted $CD3^+CD4^+$ or $CD3^+CD8^+$ T cells from donors at the time of allograft donation and from recipients at 6 months, 1 year, and 3 years posttransplant. Clonality of the *TRA* and *TRB* repertoires in individual samples was highly correlated (Pearson correlation coefficient, $r = 0.973$, $p = 1.047 \times 10^{-5}$), supporting the use of only *TRB* sequencing for subsequent studies.

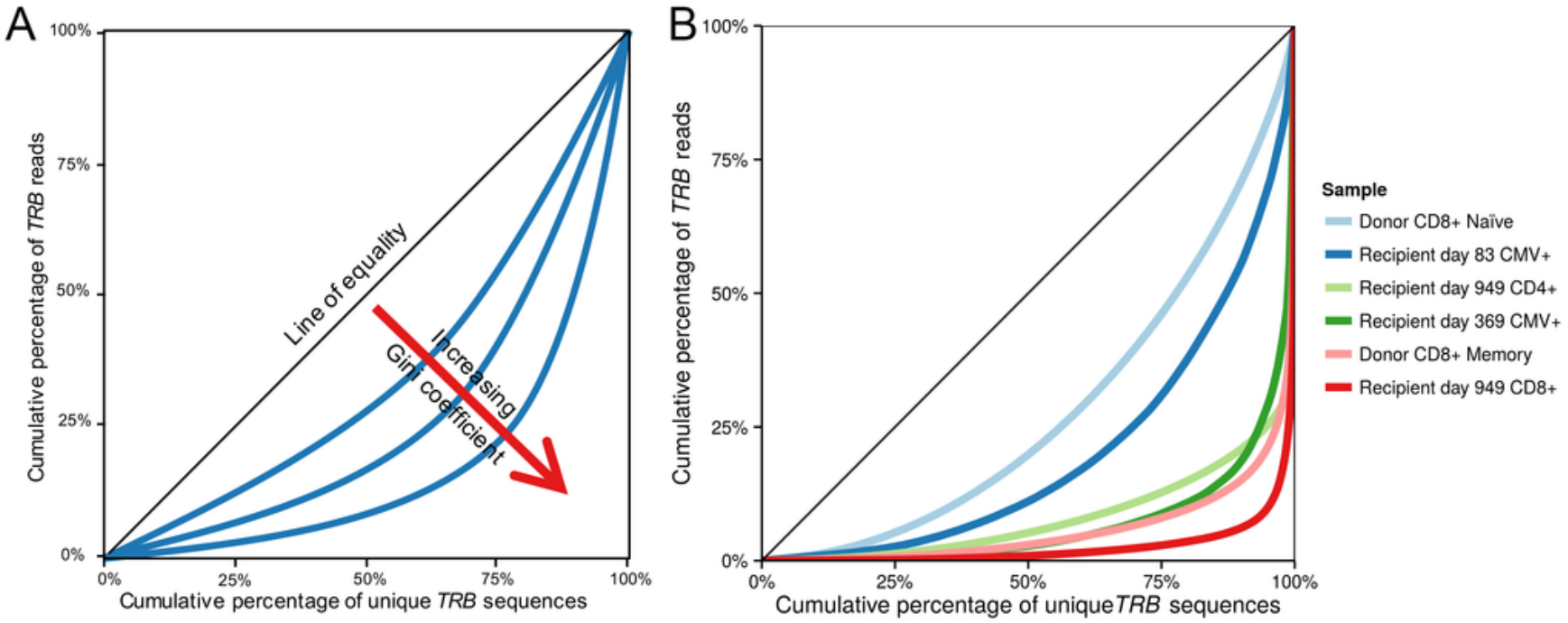

Figure S4. Lorenz curves for *TRB* repertoires in blood samples from recipient 002-037 and his donor. [A] The Lorenz curve may be used to graphically represent *TRB* clonality, wherein the x-axis represents the cumulative percentage of unique *TRB* sequences and the y-axis represents the cumulative percentage of *TRB* reads. A line passing through the origin with a slope of 1 reflects equal frequencies of all clones. The Gini coefficient is the ratio of the area between the line of equality and the observed Lorenz curve over the total area under the line of equality. The higher the coefficient, the more unequal the distribution (i.e. the more clonal the T-cell population). **[B]** Lorenz curve demonstrating increasing Gini coefficient (i.e. greater clonality) in CMV-tetramer-sorted peripheral blood T cells at later timepoints posttransplant in patient 002-037. Also apparent is a higher Gini coefficient for 002-037 donor $CD8^+$ memory T cells than for $CD8^+$ naïve T cells.

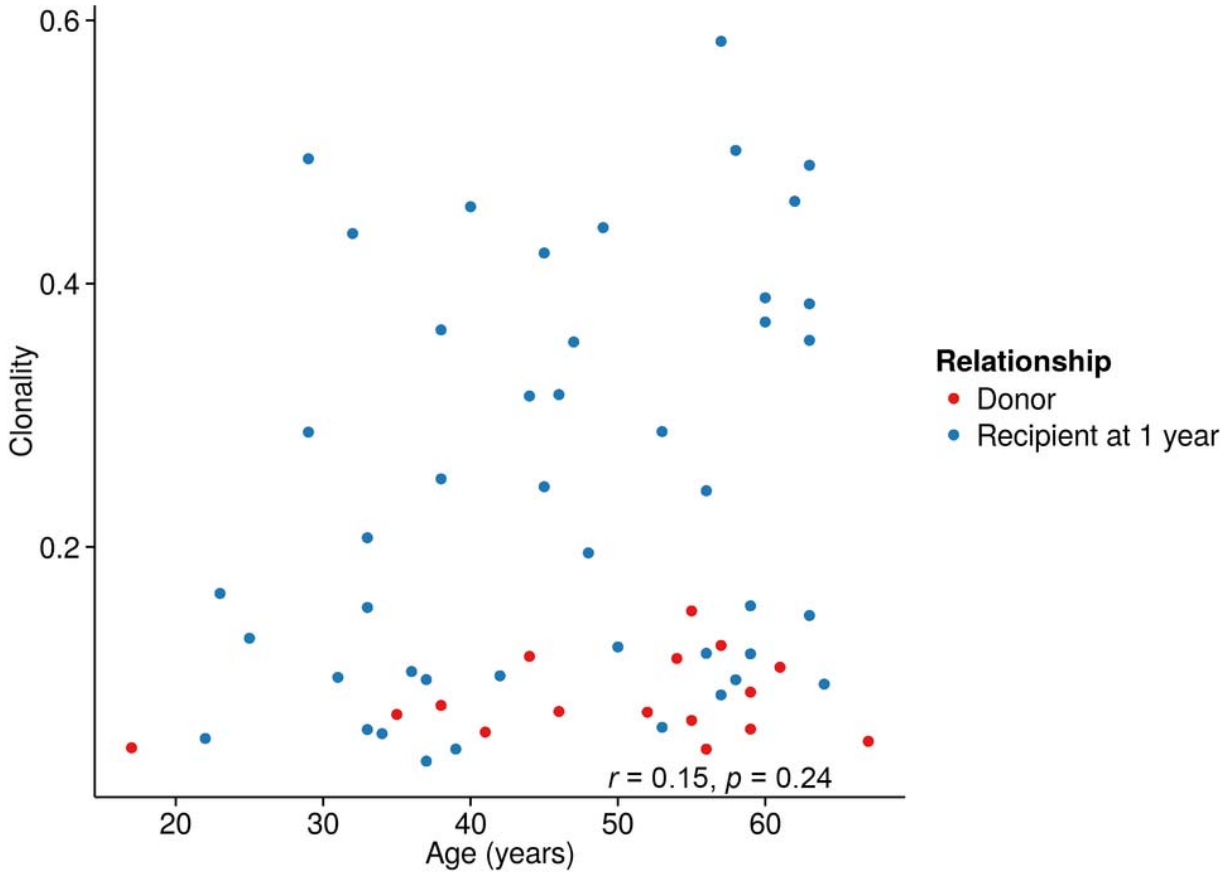

Figure S5. No correlation between donor or recipient age and *TRB* repertoire clonality. Unsorted peripheral blood samples from donors or from recipients at 1-year posttransplant were assessed by *TRB* sequencing. The age of either donor or recipient was not significantly correlated with clonality of the *TRB* repertoire (Pearson correlation coefficient, $r = 0.153$, $p = 0.238$).

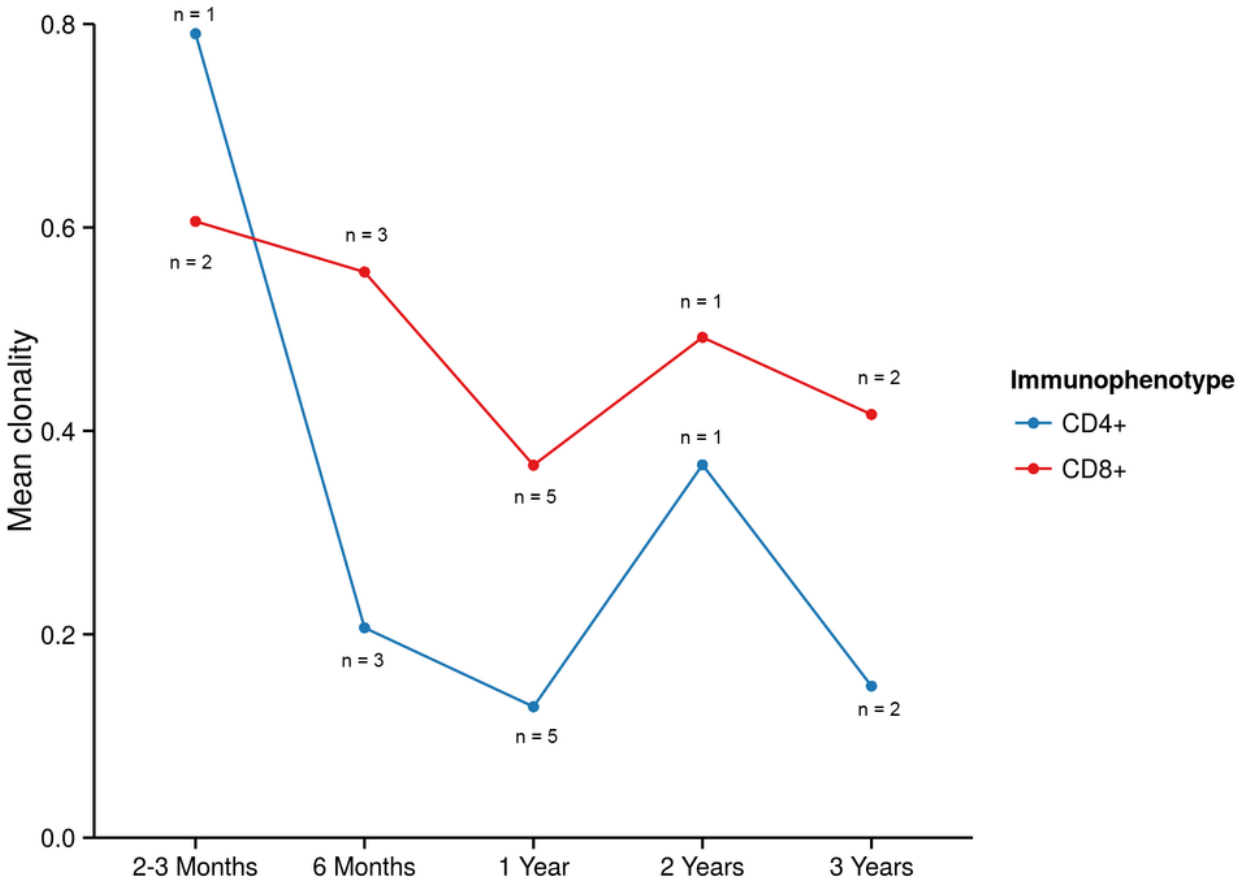

Figure S6. Clonality of blood CD8$^+$ *TRB* repertoires is higher than that of blood CD4$^+$ repertoires. Mean *TRB* clonality was consistently higher for CD4$^+$ T cells compared with CD8$^+$ T cells in the peripheral blood after 2-3 months posttransplant.

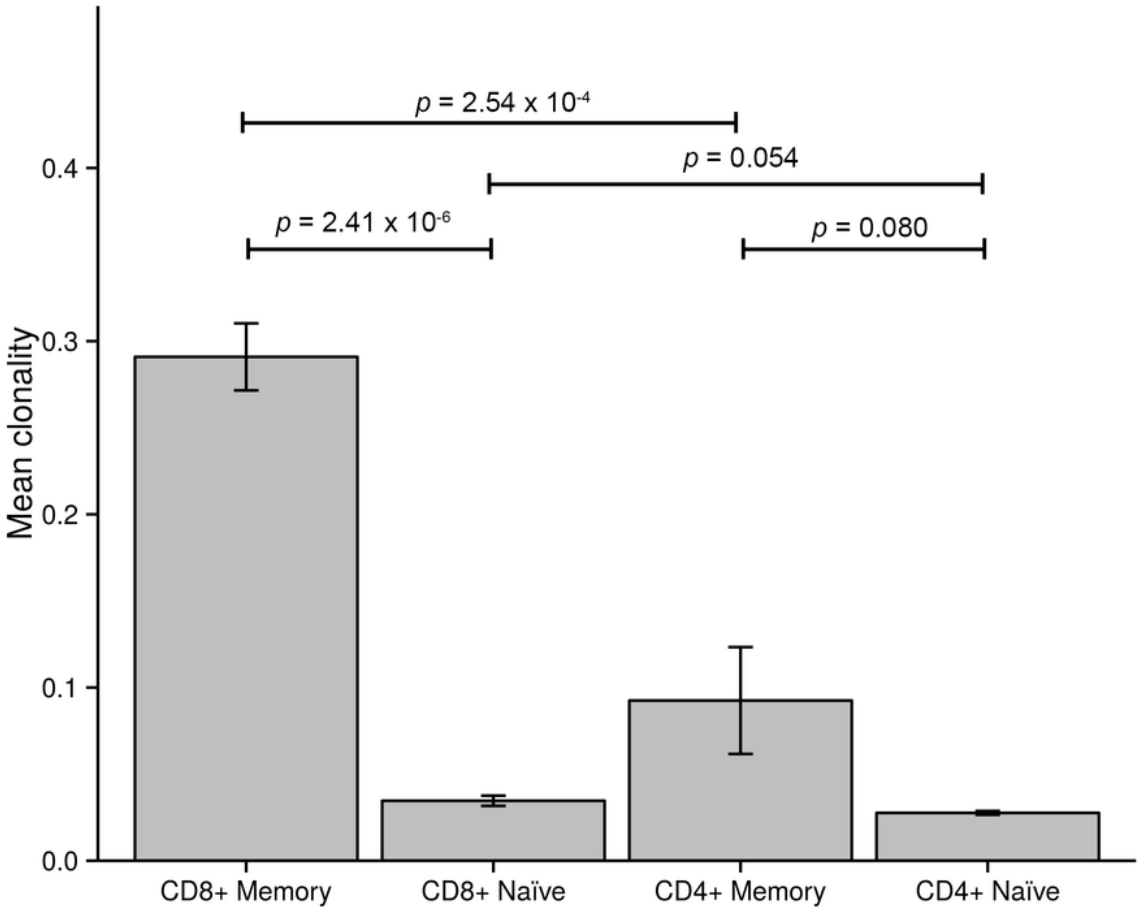

Figure S7. *TRB* clonality is higher among memory T cells compared with naïve T cells and among memory CD8$^+$ T cells compared with memory CD4$^+$ T cells. Mean *TRB* clonality was evaluated in CD8$^+$ memory, CD8$^+$ naïve, CD4$^+$ memory, and CD4$^+$ naïve T cells sorted from 8 recipient and donor peripheral blood samples obtained between 1 month and 5 years posttransplant (recipients) or at the time of allograft donation (donors). The clonality of CD8$^+$ *TRB* repertoires was higher than that of CD4$^+$ repertoires, and the clonality of memory repertoires was higher than that of naïve repertoires.

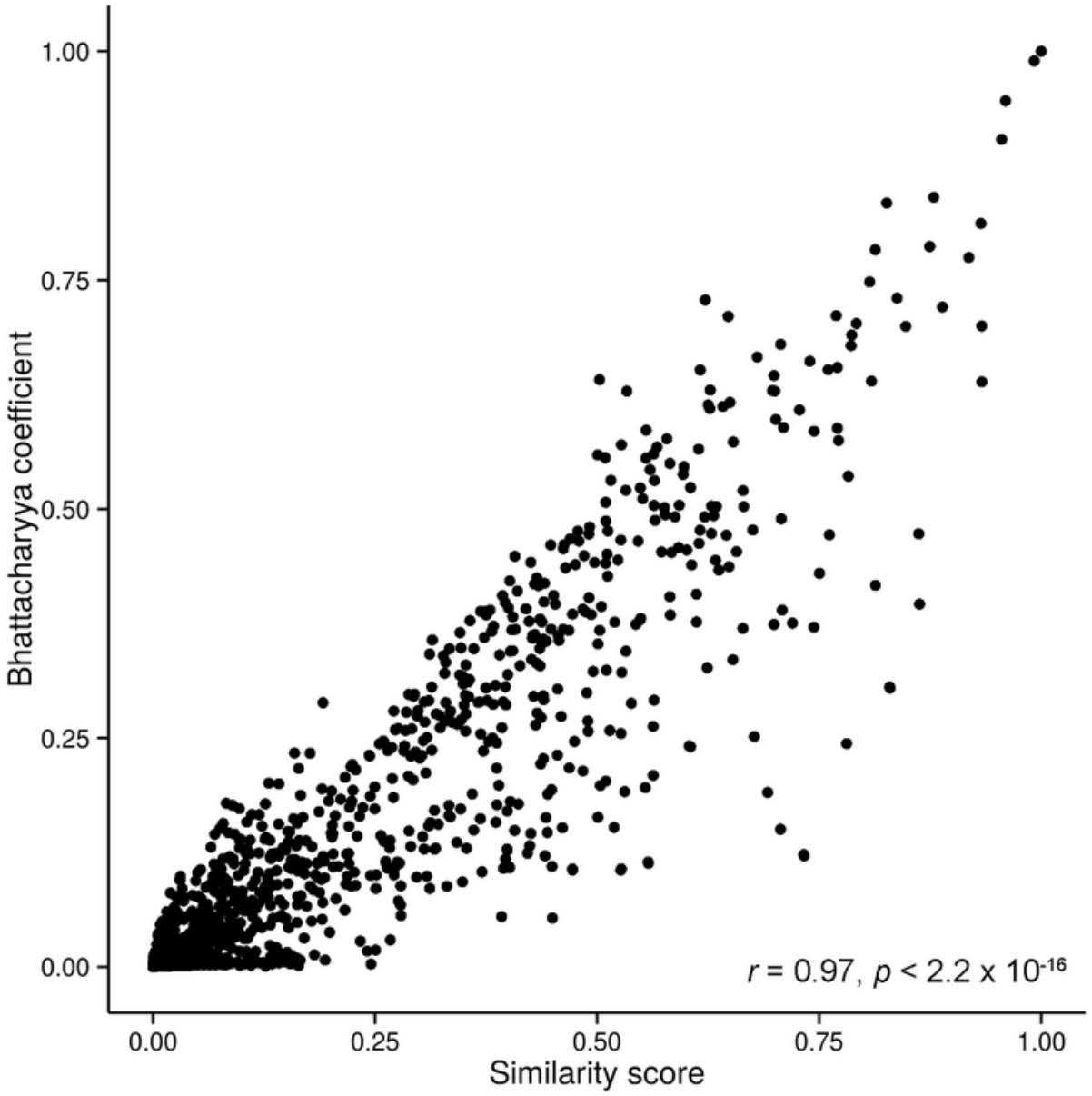

Figure S8. The Bhattacharyya coefficient and similarity score are highly correlated. Bhattacharyya coefficients and similarity scores were calculated for all pairwise comparisons of distinct *TRB* repertoires from patients or their bone marrow donors. These two metrics of repertoire similarity were highly correlated ($r = 0.97$, $p < 2.2 \times 10^{-16}$).

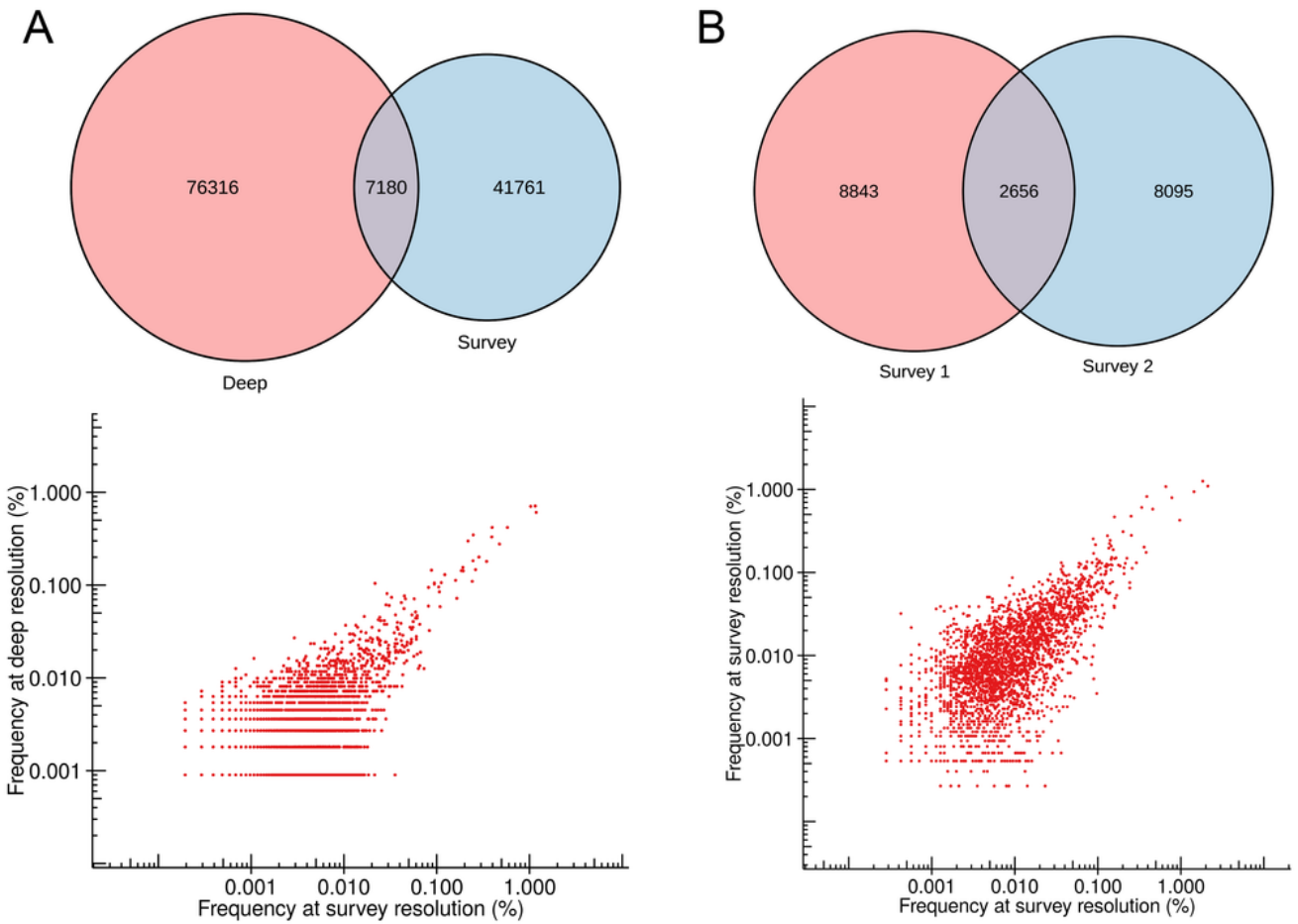

Figure S9. Clonality assessment of the *TRB* repertoire is robust to sequencing depth and reproducible. [A] Genomic DNA isolated from a sample of peripheral blood mononuclear cells from donor 002-037 underwent *TRB* repertoire sequencing at two different sequencing depths (61,256 total reads and 109,400 total reads). The *TRB* repertoire clonality inferred from the two sequence datasets was 0.073 and 0.045, respectively. The Venn diagram (top) revealed surprisingly limited global overlap between the sets of unique *TRB* sequences observed in the two datasets, which was largely due to the fact that both datasets contain a large number of low frequency (<0.01%) unique sequences that were present in one dataset but not the other. The scatter plot (bottom), however, demonstrated that the vast majority of unique *TRB* sequences with frequency >0.01% were observed in both datasets at comparable frequencies. **[B]** Similar data are shown for two independent *TRB* sequence datasets generated at survey depth from a single genomic DNA sample extracted from PBMC from recipient 002-018. Calculated *TRB* repertoire clonality for the two samples was 0.108 and 0.105.

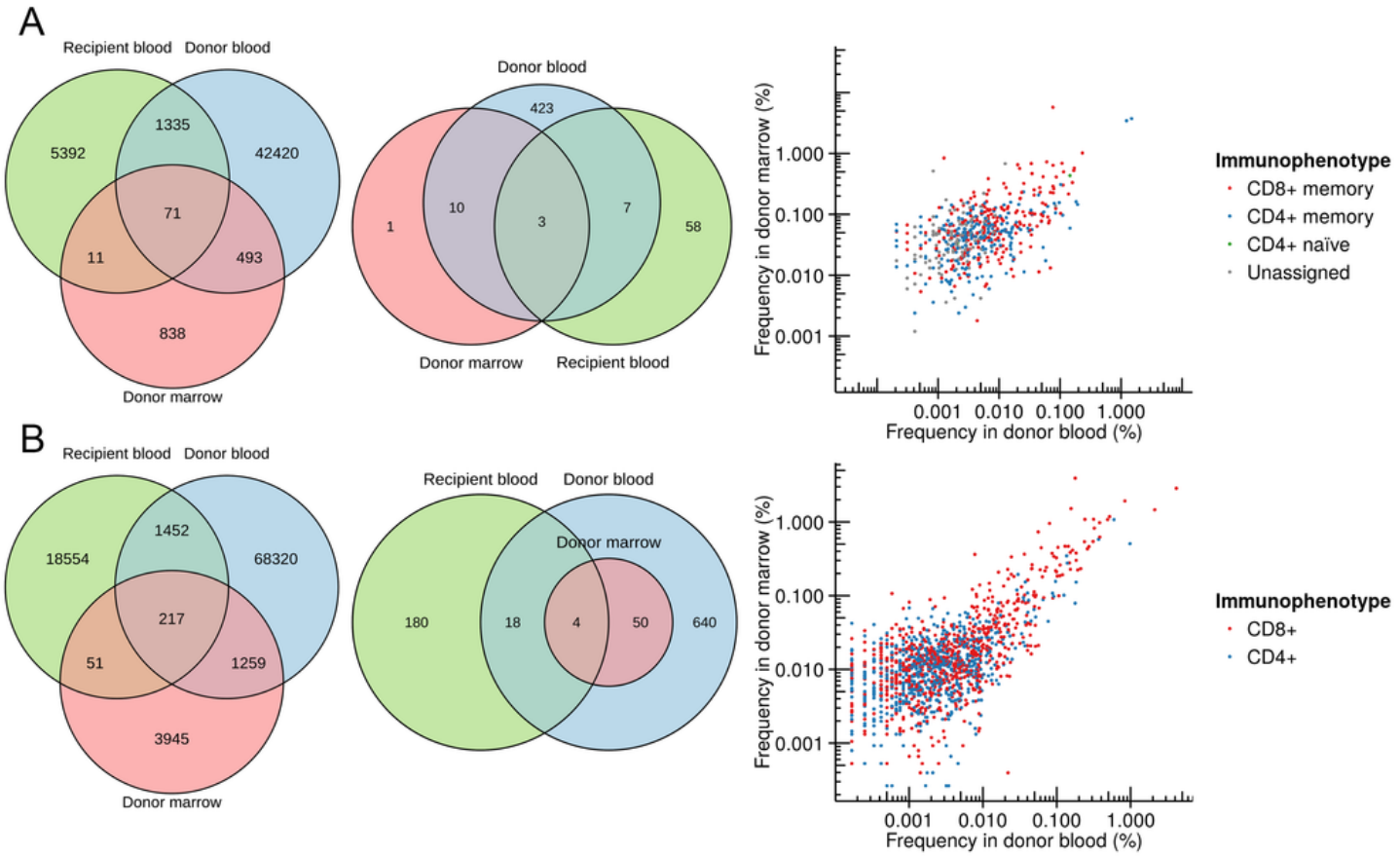

Figure S10. Global similarity and sequence sharing between donor peripheral blood and bone marrow *TRB* repertoires. In addition to donor 001-006, the data from whom are shown in Figure 6, samples of donor bone marrow were available from donors 001-007 and 001-023. Samples of blood and bone marrow from these two donors underwent flow cytometric sorting and *TRB* sequencing, and the observed repertoires were compared to *TRB* repertoires of their respective recipients. **[A]** The Venn diagram on the left shows the overlap between the *TRB* sequence repertoires observed in recipient 001-007 at 1 month and 2 months posttransplant, in donor 001-007 blood on the day of allograft donation, and in donor 001-007 bone marrow. The Venn diagram on the right compares the *TRB* repertoires of the same samples but shows only the *TRB* sequences in the top percentile of frequency. The scatter plot compares the frequency in 001-007 donor blood and donor bone marrow of *TRB* sequences that were observed in both physical compartments (n=564), and is color coded according to the immunophenotypic compartment to which each sequence was mapped. Although the global overlap was relatively low between bone marrow and peripheral blood, there was high concordance among the dominant clones present in both compartments. **[B]** Similar findings are shown for the *TRB* repertoires observed in recipient 001-023 peripheral blood at 1 month, 2 months, 6 months, and 1 year posttransplant, donor 001-023 peripheral blood on the day of allograft donation, and donor 001-023 bone marrow.

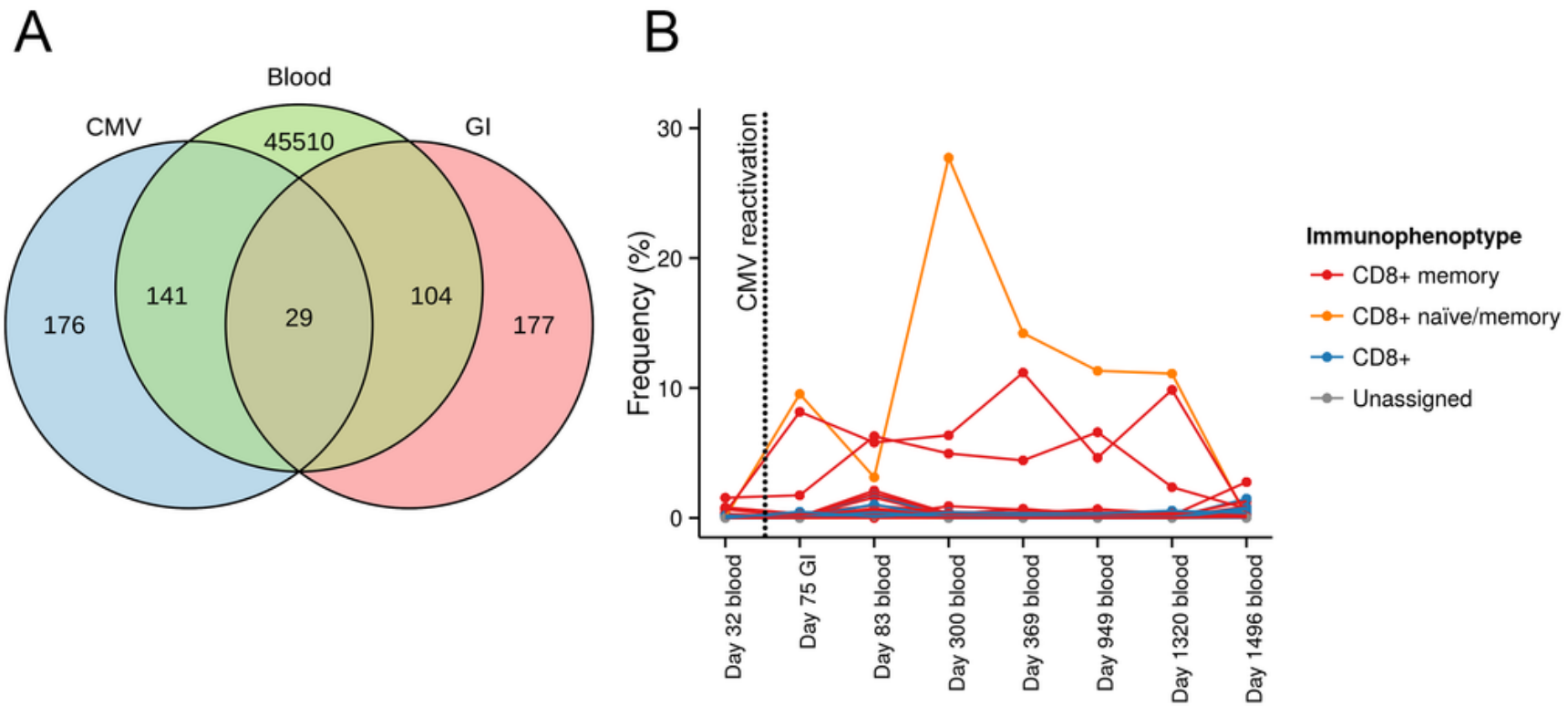

Figure S11. CMV-associated *TRB* sequences are detectable in the blood and gastrointestinal tract. In this CMV-seropositive patient who received an allograft from a CMV-seropositive donor, CMV pp65-specific tetramer^{+} (HLA-B7 and HLA-A2-restricted) T cells were sorted from peripheral blood and mapped to their frequency in the peripheral blood and gastrointestinal tissue. **[A]** Venn diagram shows the overlapping *TRB* sequences from patient 002-037 peripheral blood at 1 month to 3 years posttransplant, day +75 upper and lower gastrointestinal biopsies, and CMV. **[B]** Frequency of the 29 sequences detected in the blood, gastrointestinal biopsies, and CMV tetramer-sorted population tracked over time are shown.

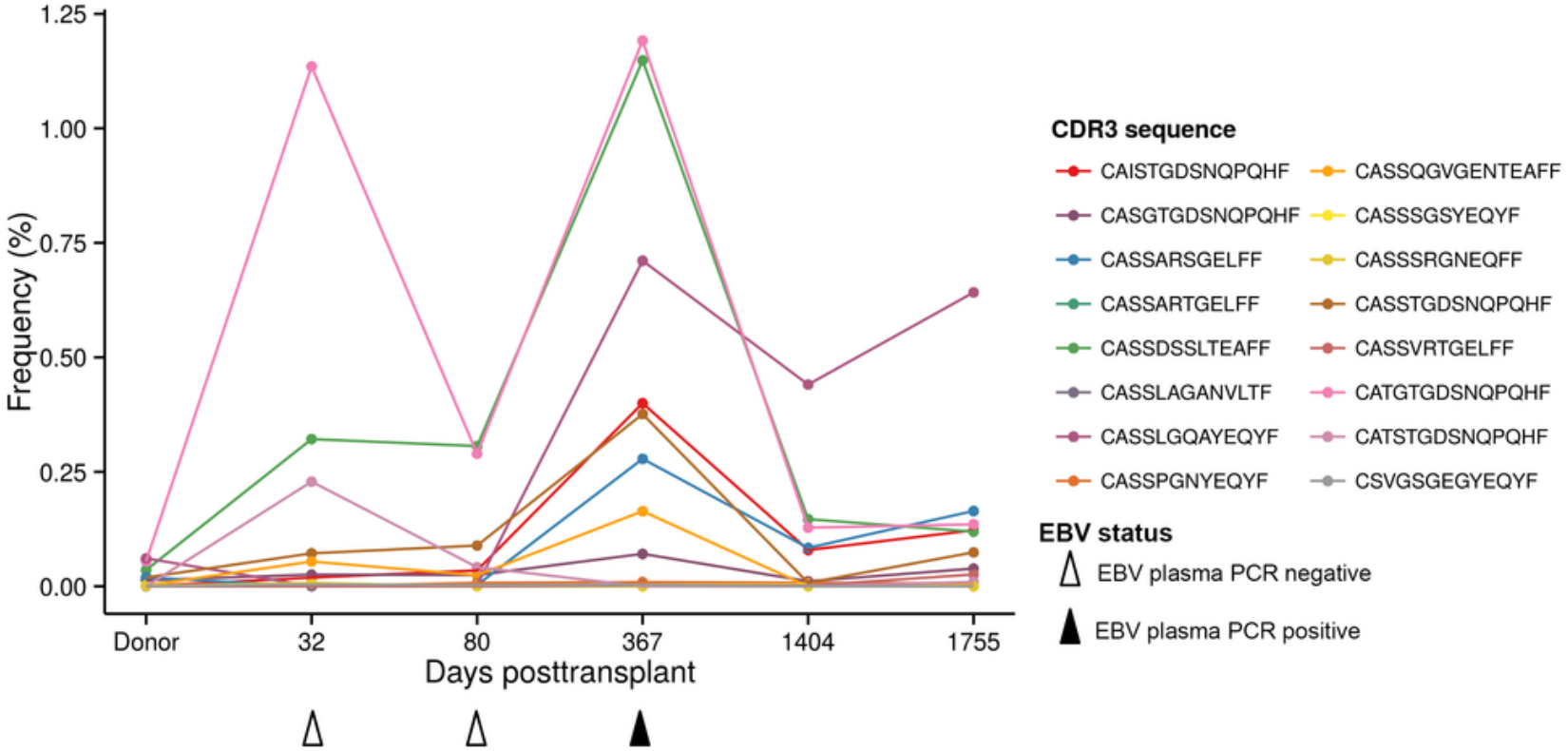

Figure S12. Public EBV-specific *TRB* sequences are detectable. Frequency of 16 EBV-associated public *TRB* sequences detectable in the peripheral blood of subject 002-014 and followed over time. Monitoring for EBV viremia was not prospectively performed during the posttransplant treatment course for patients on this clinical study, and none developed clinical signs of EBV-related disease. However, available frozen plasma samples for subject 002-014 were retrospectively assessed for EBV DNA by the clinical laboratory. The increase in frequency of many of public EBV-associated *TRB* sequences at 1-year posttransplant correlated with detection at low-level (100 copies) of EBV DNA in plasma. The patient's HLA-B genotype (HLA-B*08:01 and HLA-B*35:01) matches the MHC restriction of the $CD8^+$ EBV-specific T-cell clones that carry these public *TRB* sequences.